

**Continuity of the asymptotics of expected zeros of  
fewnomials**

by

Timothy Tran

A dissertation submitted to The Johns Hopkins University in conformity with the  
requirements for the degree of Doctor of Philosophy.

Baltimore, Maryland

March, 2014

© Timothy Tran 2014

All rights reserved

# Abstract

In “Random complex fewnomials, I,” the limiting formula of the (normalized) expected distribution of complex zeros of a system of  $k$  random  $m$ -variate fewnomials where the coefficients are taken from the  $SU(m + 1)$  ensemble and the spectra are chosen uniformly at random is determined. We recall this result and show the limiting formula is a  $(k, k)$ -form with continuous coefficients.

# Acknowledgments

I would first like to thank my advisor, Dr. Bernard Shiffman, for his guidance and wisdom throughout my graduate studies. I would next like to thank my friend and colleague, Arash Karami, with whom I could freely ask questions, discuss mathematics, and share my ideas. Finally, I would like to thank all my friends for their support, especially Josh Kaufmann, John and Lauren Ross, and Vitaly Lorman.

# Dedication

I dedicate this dissertation to my parents, Toan Tran and Ann Cao, and my sister, Crystal Tran. My deepest thanks for all your love and support.

# Contents

<b>Abstract</b>	<b>ii</b>
<b>Acknowledgments</b>	<b>iii</b>
<b>List of Figures</b>	<b>vii</b>
<b>1 Introduction</b>	<b>1</b>
1.1 Past results about fewnomials . . . . .	1
1.2 Probabilistic results . . . . .	4
<b>2 Background</b>	<b>11</b>
2.1 Main background . . . . .	12
2.2 Rewriting the function $F_n$ . . . . .	15
<b>3 Notation, Functions, and Motivation</b>	<b>18</b>
3.1 Notation and the function $b$ . . . . .	19
3.2 Motivation . . . . .	23

## CONTENTS

3.3	Defining the functions $\Lambda$ , $B$ , and $h$	29
3.4	Properties of the function $B$	32
3.4.1	Along $s = 0$ .	32
3.4.2	Along $s = 1$ .	35
3.4.3	Summary for $B$	38
3.5	Properties of the function $h$	40
3.5.1	Along $(0, \rho, x)$ .	40
3.5.2	Along $(b(x, \rho), \rho, x)$ .	42
3.5.3	Summary for $h$	44
<b>4</b>	<b>Proof of the Theorem</b>	<b>46</b>
4.1	Rewriting $D(t, \rho)$ as an integral	47
4.2	The first partial derivative of $G$	53
4.3	The second partial derivative of $G$	54
4.3.1	The term $-\int_0^\infty n[1 - D]^{n-1}\tilde{S}dt$ is continuously differentiable	55
4.3.2	The term $-\int_0^\infty n[1 - D]^{n-1}Sdt$ is continuously differentiable	56
4.3.3	Summary for $G$	63
4.4	Conclusion of the proof	64
	<b>Bibliography</b>	<b>65</b>
	<b>Vita</b>	<b>68</b>

# List of Figures

1.2.1 Density functions $\frac{\partial^2}{\partial \rho^2} F_n$ for $n = 2$ ( $\bullet$ ), $n = 4$ ( $\blacksquare$ ), $n = 8$ ( $\blacklozenge$ ), $n = 32$ ( $\blacktriangle$ ), and $n = 128$ ( $\blacktriangledown$ ). Also the density of $\omega_{\text{FS}}$ (---).	7
1.2.2 Quotient of density functions $\frac{\partial^2}{\partial \rho^2} F_n$ over the density of $\omega_{\text{FS}}$ for $n = 2$ ( $\circ$ ), $n = 4$ ( $\square$ ), $n = 8$ ( $\lozenge$ ), $n = 32$ ( $\triangle$ ), $n = 128$ ( $\triangledown$ ).	8
1.2.3 Normalized density functions $\frac{1}{\mathbb{E}_n(X)} \frac{\partial^2}{\partial \rho^2} F_n$ for $n = 2$ ( $\circ$ ), $n = 4$ ( $\square$ ), $n = 8$ ( $\lozenge$ ), $n = 32$ ( $\triangle$ ), and $n = 128$ ( $\triangledown$ ). Also the density of $\omega_{\text{FS}}$ (---).	10
3.2.1 The functions $b(\lambda, \rho)$ and its two inverse branches $g_1(t, \rho)$ and $g_{-1}(t, \rho)$ for $\rho = 0.5$ .	24
3.2.2 A 3D plot and a 10-level contour plot of $b(\lambda, \rho)$ for $\rho = (0.5, 0)$ .	26
3.2.3 The ray with vertex $\mu(\rho_0)$ and angle $\theta$ intersects the curve $b(\lambda, \rho_0) = t_0$ at the point $a$ . Equivalently, there is a point $x$ on the boundary such that the line segment from $\mu(\rho)$ to a point $x$ also intersects the curve at $a$ .	27

# Chapter 1

## Introduction

Concerned with the distribution of complex zeros of a system of random complex fewnomials, we begin this work by covering some older results on fewnomials, cite two theorems which determine the distribution of zeros from a probabilistic point of view, and follow up with the statement of our result.

Before continuing, we also note that Sections [1.1](#) and [2.1](#) are intended to supplement our result, but are inessential with respect to our proof.

### 1.1 Past results about fewnomials

A. G. Kushnirenko conjectured in the late 1970s that the number of real roots of a general system of  $m$  polynomials in  $m$  variables can be bounded above by a function of the number of nonzero terms which appear [[Kho91](#)]. Equivalently, the conjecture



## CHAPTER 1. INTRODUCTION

states that the complexity of a polynomial system can be measured by the number of nonzero terms among the polynomials in the system instead of the degrees of the polynomials in the system.

For the case when  $m = 1$ , an upper bound for the number of positive roots of a univariate polynomial can be found using Descartes' Rule of Signs [Des54]. Ordering the terms of a polynomial by increasing exponent, the rule states that the number of positive roots is equal to the number of sign changes between consecutive non-zero coefficients or less than it by a multiple of two. As a corollary, if a polynomial had  $n$  nonzero coefficients, it would have at most  $n - 1$  positive roots, because it can have at most  $n - 1$  sign changes. For example, for any real constants  $a$  and  $b$ , the polynomial  $x^{1000} + ax + b$  would have at most two positive roots.

In general, A. Khovanskii's work on fewnomial systems in the 1980s determines an upper bound for zeros in the positive orthant of  $\mathbb{R}^m$  which only depend on the number of nonzero coefficients and the dimension  $m$  [Kho91]. Unfortunately, his upper bound is not sharp. More on his result is detailed below.

Given a system  $P$ , let  $\Delta_j$  denote the Newton polytope of  $P_j$ , i.e., the convex hull of the exponents appearing nontrivially in  $P_j$ . Then let  $V(\Delta_1, \dots, \Delta_m)$  denote the mixed volume of  $\Delta_1, \dots, \Delta_m$ . By the Bernstein-Kushnirenko Theorem, the total number of solutions to a general system  $P$  is  $m!V(\Delta_1, \dots, \Delta_m)$  [Ber75, KK12]. Next consider the real  $m$ -torus  $\mathbb{T}^m = (\mathbb{R}/[0, 2\pi])^m$ . Then  $(\mathbb{C} \setminus \{0\})^m$  can be identified with  $\mathbb{R}^m \times \mathbb{T}^m$  under the map  $z = [\exp(\frac{1}{2}\rho_1 + i\theta_1), \dots, \exp(\frac{1}{2}\rho_m + i\theta_m)] \mapsto (\rho, \theta)$  and we

## CHAPTER 1. INTRODUCTION

call  $\theta$  the argument of  $z$ . Fixing an angular sector  $U \subset \mathbb{T}^m$ , let  $N(P, U)$  denote the number of zeros of  $P$  with arguments lying in  $U$  and let

$$S(P, U) = m! V(\Delta_1, \dots, \Delta_m) \frac{\text{Vol}(U)}{\text{Vol}(\mathbb{T}^m)}.$$

Note that the quantity  $S(P, U)$  may be interpreted as the expected number of zeros of  $P$  lying in  $U$  if the zeros were uniformly distributed.

Next, let  $\Delta^* \subset \mathbb{R}^m$  denote the intersection of the sets  $\{y \in \mathbb{R}^m \mid |\langle \alpha, y \rangle| < \pi\}$  as  $\alpha$  ranges over the exponents appearing in the spectra  $\Delta_i$ . Then let  $\overline{\Delta}^*$  denote the image of the region  $\Delta^*$  under the quotient homomorphism  $\phi : \mathbb{R}^m \rightarrow \mathbb{T}^m$  and  $\Pi(\overline{\Delta}^*, U)$  be the minimal number of translates of  $\overline{\Delta}^*$  needed to cover the boundary of  $U$ . Khovanskii's result on real fewnomial systems says the following:

**Theorem 1** (Khovanskii, Theorem 2 [Kho91]). *There exists an explicit function  $\varphi(n, m)$  of two natural number variables  $(n, m)$  such that for each system of polynomial equations  $P_1 = \dots = P_m = 0$  with Newton polyhedra  $\Delta_1, \dots, \Delta_m$  that contains at most  $n$  monomials, nonsingular at infinity in the region  $\arg z \in U$  of the space  $(\mathbb{C} \setminus \{0\})^m$ , the following relation holds:*

$$|N(P, U) - S(P, U)| < \varphi(n, m) \Pi(\overline{\Delta}^*, U). \quad (1.1.1)$$

Taking a sequence of balls  $U_j \subset \mathbb{T}^m$  shrinking to  $\{0\} \subset \mathbb{T}^m$ , then  $\Pi(\overline{\Delta}^*, U_j) \rightarrow 1$  and  $S(P, U_j) \rightarrow 0$ . It then follows from (1.1.1) that

$$|N_{\mathbb{R}}(P)| \leq \varphi(n, m), \quad (1.1.2)$$

## CHAPTER 1. INTRODUCTION

an upper bound for the number of real zeros of a polynomial system  $P$  which depends only on the number of nonzero terms  $n$  and the number of variables  $m$ .

Explicitly, Khovanskii finds the number of nondegenerate roots of a polynomial system  $P = (P_1 = \cdots = P_m = 0)$  lying in the positive orthant of  $\mathbb{R}^m$  is at most

$$2^{n(n-1)/2}(m+1)^n,$$

where  $n$  is the number of distinct monomials that appear with nonzero coefficient in at least one of the polynomials [Kho91]. However, this bound is not sharp, and even an improvement of the bound to

$$\frac{e^2 + 3}{4} 2^{l(l-1)/2} m^l,$$

where  $m + l + 1$  is the number of monomials [BS07] is only asymptotically sharp (for fixed  $l$  and large  $m$ ) [BRS08]. For related results, see also [BBS07, BS11].

Of course, even if a sharp bound is determined, the bound is unlikely to represent the typical distribution of zeros to complex fewnomial systems. Considering the typical behavior of the zero distribution prompts studying complex fewnomial systems probabilistically.

## 1.2 Probabilistic results

About ready to state our own result, we first cite two theorems that help put our own theorem into perspective; pertinent background leading up to these theorems

## CHAPTER 1. INTRODUCTION

may be found in Chapter 2.

For  $z \in (\mathbb{C} \setminus \{0\})^m$ , we write  $z = [\exp(\frac{1}{2}\rho_1 + i\theta_1), \dots, \exp(\frac{1}{2}\rho_m + i\theta_m)]$ . Then we note that while Khovanskii's theorem (Theorem 1) is concerned with the angular  $(\theta_1, \dots, \theta_m)$  distribution of zeros, the results below prescribe a radial  $(\rho_1, \dots, \rho_m)$  distribution of zeros.

The first probabilistic result considers dilation of a fixed spectrum:

**Theorem 2** (Shiffman and Zelditch, Theorem 1.2 [SZ11]). *Let  $S = \{\lambda^1, \dots, \lambda^n\}$  be a fixed spectrum consisting of  $n$  lattice points in  $p\Delta$  where  $p \in \mathbb{Z}^+$ . For random  $m$ -tuples  $(P_1^N, \dots, P_m^N)$  of  $n$ -nomials in  $\text{Poly}(NS)$  of degree  $Np$ , with coefficients chosen from the  $\text{SU}(m+1)$  ensemble, the expected distribution of zeros in  $(\mathbb{C} \setminus \{0\})^m$  has the asymptotics*

$$N^{-m} \mathbb{E}_{Np|NS}[Z_{P_1^N, \dots, P_m^N}] \rightarrow p^m \left( \frac{i}{2\pi} \partial \bar{\partial} \max_{\lambda \in S} [\langle \rho, \lambda \rangle - \langle \hat{\lambda}, \log \hat{\lambda}^p \rangle] \right)^k.$$

Here,  $\hat{\lambda}^p = (p - |\lambda|, \lambda_1, \dots, \lambda_m)$  and  $\log \hat{\lambda}^p = (\log(p - |\lambda|), \log \lambda_1, \dots, \log \lambda_m)$ .

Looking at the potential function  $\rho \mapsto \max_{\lambda \in S} [\langle \rho, \lambda \rangle - \langle \hat{\lambda}, \log \hat{\lambda}^p \rangle]$  of the limit, we note that it is a piecewise linear function on  $\mathbb{R}^n$ . Thus, the limit is a singular measure supported along where the functions meet. For the case when  $k = m$ , the measure is supported along the 0-dimensional corner set.

The second probabilistic result considers choosing the spectra uniformly at random:

## CHAPTER 1. INTRODUCTION

**Theorem 3** (Shiffman and Zelditch, Theorem 1.5 [SZ11]). *Let  $1 \leq k \leq m$ , and let  $(P_1^N, \dots, P_k^N)$  be a random system of  $n$ -nomials of degree  $N$ , where the spectra  $S_j$  are chosen uniformly at random from the simplex  $N\Delta$  and the coefficients are chosen from the  $SU(m+1)$  ensemble. Then the expected zero current in  $(\mathbb{C} \setminus \{0\})^m$  has the asymptotics*

$$N^{-k} \mathbb{E}_{N,n}[Z_{P_1^N, \dots, P_k^N}] \rightarrow \left( \frac{i}{2\pi} \partial \bar{\partial} \int_{\Delta^n} \max_{j=1, \dots, n} [\langle \rho, \lambda^j \rangle - \langle \hat{\lambda}^j, \log \hat{\lambda}^j \rangle] d\lambda^1 \cdots d\lambda^n \right)^k.$$

Here,  $\hat{\lambda} = (1 - |\lambda|, \lambda_1, \dots, \lambda_m)$ ,  $\log \hat{\lambda} = (\log(1 - |\lambda|), \log \lambda_1, \dots, \log \lambda_m)$ , and  $d\lambda = m! d\lambda_1 \cdots d\lambda_m$ .

Comparing the second result with the first result, we note the appearance of the piecewise linear function  $(\rho, \lambda) \mapsto \max_{j=1, \dots, n} [\langle \rho, \lambda^j \rangle - \langle \hat{\lambda}^j, \log \hat{\lambda}^j \rangle]$  in the integrand. Paired with the earlier observation that the limit when dilating a fixed spectrum is a singular measure, it becomes unclear whether or not the limit when choosing the spectra uniformly at random is also a singular measure.

Answering the uncertainty, our work shows that the limit when choosing the spectra uniformly at random is a form with continuous coefficients. Explicitly, we have determined the following:

**Theorem 4.** *Define the function*

$$F_n : \mathbb{R}^m \rightarrow \mathbb{R}$$

$$\rho \mapsto \int_{\Delta^n} \max_{j=1, \dots, n} [\langle \rho, \lambda^j \rangle - \langle \hat{\lambda}^j, \log \hat{\lambda}^j \rangle] d\lambda^1 \cdots d\lambda^n.$$

## CHAPTER 1. INTRODUCTION

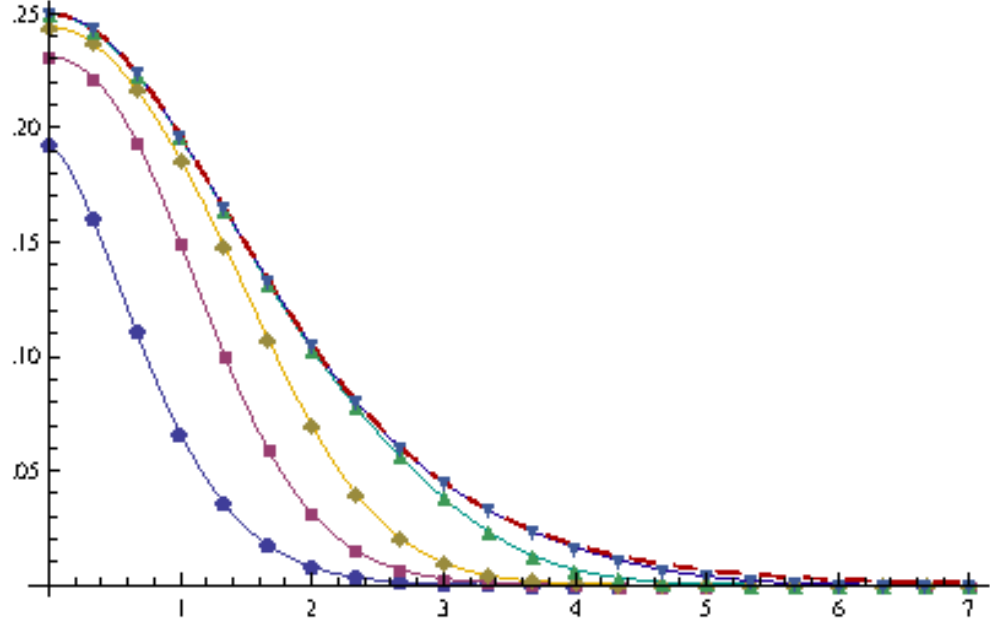


Figure 1.2.1: Density functions  $\frac{\partial^2}{\partial \rho^2} F_n$  for  $n = 2$  (blue circles),  $n = 4$  (purple squares),  $n = 8$  (yellow diamonds),  $n = 32$  (green triangles), and  $n = 128$  (dark blue inverted triangles). Also the density of  $\omega_{FS}$  (dashed red line).

Then  $F_n$  is  $\mathcal{C}^2(\mathbb{R}^m)$  and thus the distribution

$$\left( \frac{i}{2\pi} \partial \bar{\partial} \int_{\Delta^n} \max_{j=1, \dots, n} \left[ \langle \rho, \lambda^j \rangle - \langle \hat{\lambda}^j, \log \hat{\lambda}^j \rangle \right] d\lambda^1 \cdots d\lambda^n \right)^k \quad (1.2.1)$$

is a  $(k, k)$ -form with continuous coefficients.

In particular, when  $k = m$ , the distribution in (1.2.1) is given by

$$m! \det \left( \frac{1}{2\pi} \frac{\partial^2}{\partial \rho_i \partial \rho_j} F_n \right) d\rho d\theta,$$

which has a continuous density with respect to the Lebesgue measure on  $(\mathbb{C} \setminus \{0\})^m$ .

As a consequence of our main theorem, we are able to graph the density functions for values of  $n$  (for dimension  $m = 1$ ). In Figure 1.2.1, we plot the density functions

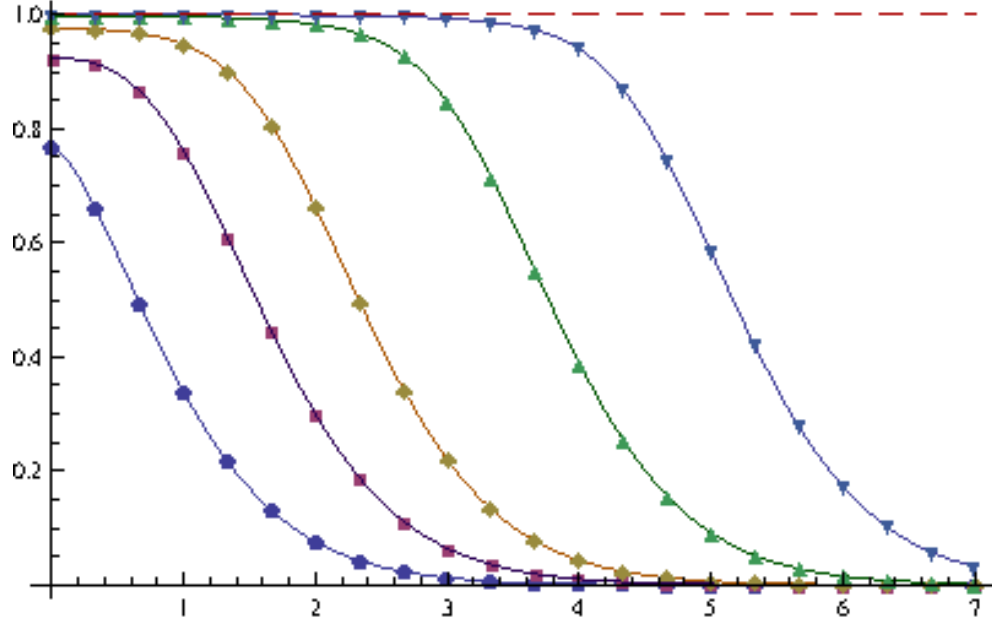


Figure 1.2.2: Quotient of density functions  $\frac{\partial^2}{\partial \rho^2} F_n$  over the density of  $\omega_{\text{FS}}$  for  $n = 2$  ( $\bullet$ ),  $n = 4$  ( $\blacksquare$ ),  $n = 8$  ( $\blacklozenge$ ),  $n = 32$  ( $\blacktriangle$ ),  $n = 128$  ( $\blacktriangledown$ ).

for  $n = 2, 4, 8, 32, 128$ . We plot them alongside the density function for the Fubini-Study form, because it is the invariant measure on projective space and the limiting distribution for the full system [SZ99]. For a slightly different point of view, we take the quotient of the density functions (for  $n = 2, 4, 8, 32, 128$ ) over the density function for the Fubini-Study form  $[(1 + e^\rho)(1 + e^{-\rho})]^{-1}$  and obtain Figure 1.2.2.

Finally, we can consider the normalized distribution with density  $[\mathbb{E}_n(X)]^{-1} \frac{\partial}{\partial \rho^2} F_n$ , where  $X$  is the integer-valued random variable on  $\text{Poly}(N)$  which maps a polynomial  $P \in \text{Poly}(N)$  to the number of nonzero roots of  $P$ ,  $\mathbb{E}_n(X) = \lim_{N \rightarrow \infty} \frac{1}{N} \mathbb{E}_{N,n}(X)$ , and  $\mathbb{E}_{N,n}(X)$  is the expected number of nonzero roots of an  $n$ -nomial of degree  $N$  where the spectra are chosen uniformly at random. Using combinatorics, it can be shown

## CHAPTER 1. INTRODUCTION

that

$$\mathbb{E}_{N,n}(X) = (N+2)\frac{n-1}{n+1}.$$

Explicitly, if the spectrum of a polynomial  $P \in \text{Poly}(N)$  is written  $S_P = \{s_1, \dots, s_n\}$  where  $s_i < s_j$  for  $i < j$ , then  $X(P) = s_n - s_1$ . Thus,  $X$  attains value between  $n-1$  and  $N$ . We also note that there are  $\binom{N+1}{n}$  possible spectra of size  $n$  for polynomials of degree  $N$ . Now we fix an integer  $k$  between  $n-1$  and  $N$  and count the number of possible spectra with  $n$  elements such that  $s_n - s_1 = k$ . Having fixed  $k$ , we deduce there are exactly  $N - k + 1$  choices for  $s_1$  (or equivalently  $s_n$ ) and we obtain

$$\begin{aligned} \mathbb{E}(X) &= \sum_{k=n-1}^N \mathbb{E}(X = k) = \sum_{k=n-1}^N k \mathcal{P}(X = k) \\ &= \sum_{k=n-1}^N k \left[ (N - k + 1) \binom{k-1}{n-2} \binom{N+1}{n}^{-1} \right]. \end{aligned}$$

Manipulating the above expression, we have

$$\begin{aligned} \mathbb{E}(X) &= \binom{N+1}{n}^{-1} \left[ \sum_{k=n-1}^N (N+2)k \binom{k-1}{n-2} - \sum_{k=n-1}^N (k+1)k \binom{k-1}{n-2} \right] \\ &= \binom{N+1}{n}^{-1} \left[ (N+2)(n-1) \sum_{k=n-1}^N \binom{k}{n-1} - n(n-1) \sum_{k=n-1}^N \binom{k+1}{n} \right] \\ &= \binom{N+1}{n}^{-1} \left[ (N+2)(n-1) \binom{N+1}{n} - n(n-1) \binom{N+2}{n+1} \right] \\ &= (N+2) \frac{n-1}{n+1} \end{aligned}$$

as desired, where the third equality comes from reindexing and applying the identity

$$\sum_{a=b}^c \binom{a}{b} = \binom{c+1}{b+1}.$$



## CHAPTER 1. INTRODUCTION

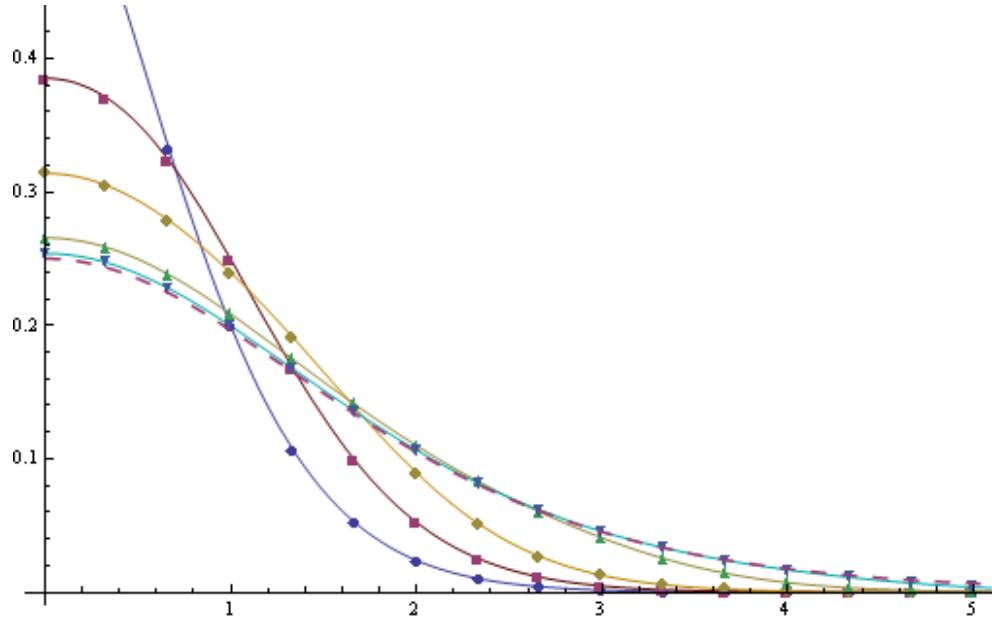


Figure 1.2.3: Normalized density functions  $\frac{1}{\mathbb{E}_n(X)} \frac{\partial^2}{\partial \rho^2} F_n$  for  $n = 2$  (—),  $n = 4$  (■),  $n = 8$  (◆),  $n = 32$  (▲), and  $n = 128$  (▼). Also the density of  $\omega_{\text{FS}}$  (---).

Note that we also have

$$\int_{\mathbb{C}} \mathbb{E}_{N,n}[Z_{P^N}] = \mathbb{E}_{N,n}(X).$$

Dividing both sides by  $N$  and taking the limit as  $N$  tends to infinity, we obtain

$$\int_{\mathbb{C}} \frac{\partial^2}{\partial \rho^2} F_n d\rho d\theta = \mathbb{E}_n(X).$$

Thus the distribution with density  $[\mathbb{E}_n(X)]^{-1} \frac{\partial^2}{\partial \rho^2} F_n$  is a probability distribution. We graph these distributions for  $n = 2, 4, 8, 32, 128$  in Figure 1.2.3.

# Chapter 2

## Background

For completeness, this chapter begins by giving key concepts used to state and prove Theorems 2 and 3 found in Section 1.2 of Chapter 1. (See also [SZ11] or [GH94].)

Essential motivation for the material discussed consists of the following:

1. Homogeneous polynomials of degree  $N$  on  $\mathbb{C}^{m+1}$  correspond to holomorphic sections of the  $N$ th power of the hyperplane section bundle.
2. The probabilistic Poincaré-Lelong formula provides an expression for the expected zero current of a subspace  $\mathcal{S}$  of  $H^0(M, L)$  in terms of the conditional Szegő kernel  $\Pi_{\mathcal{S}}$ .

We conclude the chapter by recalling a computation which doubles as the first step in the proof of our result.

## 2.1 Main background

Given an  $m$ -variate polynomial  $P$  of degree at most  $N$ , we can always associate a homogeneous  $m + 1$ -variate polynomial  $\widehat{P}$  of degree  $N$  such that  $\widehat{P}(1, \zeta_1, \dots, \zeta_m) = P(\zeta_1, \dots, \zeta_m)$ .

Then, as mentioned at the beginning of the chapter, homogeneous polynomials of degree  $N$  on  $\mathbb{C}^{m+1}$  correspond to holomorphic sections of the line bundle  $\mathcal{O}(N)$ .

As such, we denote the space of polynomials of degree at most  $N$  by  $\text{Poly}(N)$ , identify it with  $H^0(\mathbb{CP}^m, \mathcal{O}(N))$ , and place on it the inner product  $h$  given by

$$\langle f, \bar{g} \rangle = \frac{1}{m!} \int_{\mathbb{C}^m} \frac{f(z) \overline{g(z)}}{(1 + \|z\|^2)^N} \omega_{\text{FS}}^m(z), \quad \text{for } f, g \in \text{Poly}(N),$$

where  $\omega_{\text{FS}} = \frac{i}{2\pi} \partial \bar{\partial} \log(1 + \|z\|^2)$  is the Fubini-Study Kähler form on  $\mathbb{C}^m \subset \mathbb{CP}^m$ .

Let  $\alpha = (\alpha_1, \dots, \alpha_m) \in \mathbb{N}^m$ . We denote the monomial mapping  $z$  to  $z^\alpha = z_1^{\alpha_1} \dots z_m^{\alpha_m}$  as  $\chi_\alpha$ .

Since  $h$  is invariant under the torus action, the monomials  $\chi_\alpha$  determine an orthogonal basis for  $\text{Poly}(N)$  with respect to  $h$  (true for any inner product invariant under the torus action).

Next we define  $m_\alpha$  to be the normalization of  $\chi_\alpha$  with respect to the  $h$ . Then  $m_\alpha$  form an orthonormal basis for  $\text{Poly}(N)$  with respect to  $h$ .

We now obtain the expression of a polynomial  $P \in \text{Poly}(N)$  as

$$P = \sum_{\alpha \in \mathbb{N}^m, |\alpha| \leq N} c_\alpha m_\alpha,$$

## CHAPTER 2. BACKGROUND

where  $|\alpha| = \alpha_1 + \cdots + \alpha_m$ . Now define the induced Gaussian (probability) measure  $\gamma_N$  on  $\text{Poly}(N)$  as the measure such that the coefficients  $c_\alpha$  are independent complex normal random variables. Having placed a measure on  $\text{Poly}(N)$ , we define the  $\text{SU}(m+1)$  ensemble as the spaces  $\text{Poly}(N)$  together with their corresponding measures  $\gamma_N$ .

Now fix a subspace  $\mathcal{S}$  of  $\text{Poly}(N)$ . For polynomials  $P_1, \dots, P_k \in \mathcal{S}$  and  $k \leq m$ , we consider the set of zeros  $Z_{P_1, \dots, P_k} = \{z \in (\mathbb{C} \setminus \{0\})^m \mid P_1(z) = \cdots = P_k(z) = 0\}$  and the associated current of integration  $[Z_{P_1, \dots, P_k}] \in \mathcal{D}'^{k,k}((\mathbb{C} \setminus \{0\})^m)$  given by

$$([Z_{P_1, \dots, P_k}], \psi) = \int_{Z_{P_1, \dots, P_k}} \psi, \quad \text{for } \psi \in \mathcal{D}^{m-k, m-k}((\mathbb{C} \setminus \{0\})^m).$$

Then the expected zero current, if it exists, is a  $(k, k)$ -current  $\mathbb{E}_{\mathcal{S}}[Z_{P_1, \dots, P_k}]$  such that

$$(\mathbb{E}_{\mathcal{S}}[Z_{P_1, \dots, P_k}], \psi) = \mathbb{E}_{\mathcal{S}}([Z_{P_1, \dots, P_k}], \psi), \quad \psi \in \mathcal{D}^{m-k, m-k}((\mathbb{C} \setminus \{0\})^m).$$

Speaking of expected zero currents, we recall the probabilistic Poincaré-Lelong formula which assists in their computation. It says the following [SZ99, SZ08, SZ11]:

**Proposition.** *Let  $(L, h)$  be a Hermitian line bundle on a compact Kähler manifold  $M$ . Let  $\mathcal{S}$  be a subspace of  $H^0(M, L)$  endowed with a Hermitian inner product and we let  $\gamma$  be the induced Gaussian probability measure on  $\mathcal{S}$ . Then the expected zero current of a random section  $s \in \mathcal{S}$  is given by*

$$\mathbb{E}_{\gamma}(Z_s) = \frac{\sqrt{-1}}{2\pi} \partial \bar{\partial} \log \Pi_{\mathcal{S}}(z, z) + c_1(L, h).$$

*If  $\mathcal{S}_j$  is a base-point-free linear system with Gaussian probability measure  $\gamma_j$ , for*

## CHAPTER 2. BACKGROUND

$1 \leq j \leq k$  (where  $1 \leq k \leq m$ ), then the expected value of the simultaneous zero current of  $k$  independent random sections  $s_1 \in \mathcal{S}_1, \dots, s_k \in \mathcal{S}_k$  is given by

$$\mathbb{E}_{\gamma_1, \dots, \gamma_k}(Z_{s_1, \dots, s_k}) = \bigwedge_{j=1}^k \left( \frac{\sqrt{-1}}{2\pi} \partial \bar{\partial} \log \Pi_{\mathcal{S}_j}(z, z) + c_1(L, h) \right),$$

which is a smooth form.

Applying the probabilistic Poincaré-Lelong formula, computing the expected zero currents is partially reduced to computing the conditional Szegő kernel  $\Pi_{N|S}$ .

However, note that given an orthonormal basis  $\varphi_1, \dots, \varphi_n$  for  $\mathcal{S}$ , then  $\Pi_{\mathcal{S}}(z, w) = \varphi_1(z)\overline{\varphi_1(w)} + \dots + \varphi_n(z)\overline{\varphi_n(w)}$ . Thus, computing the conditional Szegő kernel can be reduced to finding an orthonormal basis for  $\mathcal{S}$  with respect to the inner product  $\gamma$ .

Now consider the following two subsets of  $\text{Poly}(N)$  and their corresponding conditional probability measures:

1. Given  $P \in \text{Poly}(N)$ , we define the spectrum of  $P$  to be the set  $S_P = \{\alpha \in \mathbb{N}^m \mid c_\alpha \neq 0\}$ . Then choose a set of exponents  $S \subset \{\alpha \in \mathbb{N}^m \mid |\alpha| \leq N\}$  and consider the subspace  $\text{Poly}(S)$  of  $\text{Poly}(N)$  consisting of polynomials  $P$  whose spectrum  $S_P$  is contained in  $S$ . Restricting  $\gamma_N$  to this subspace, we denote the conditional probability by  $\gamma_{N|S}$ .
2. Now instead fix a natural number  $n$  and consider the set of  $n$ -nomials, polynomials  $P$  whose spectrum  $S_P$  have cardinality at most  $n$ . In this situation, we shall consider the measure  $\gamma_{N|n}$  such that the restriction of  $\gamma_{N|n}$  to  $\text{Poly}(S)$  is

## CHAPTER 2. BACKGROUND

$\frac{1}{|C(N,n)|}\gamma_{N|S}$ , where  $C(N,n)$  is the collection of sets  $S \subset \{\alpha \in \mathbb{N}^m \mid |\alpha| \leq N\}$  with cardinality  $n$  and  $|C(N,n)|$  is the cardinality of  $C(N,n)$ .

We note that the first subset is the primary concern of Theorem 2. In Theorem 2, a single spectrum  $S$  is chosen and the sequence of spectra  $\{N \cdot S\}_{N \in \mathbb{N}}$ , dilations of the spectrum  $S$ , determines a sequence of expected zero currents which is shown to have a limit when normalizing by a factor of  $N$ .

Similarly, the second subset is the primary concern of Theorem 3. Letting  $P_N$  be the set of  $n$ -nomials of degree at most  $N$ , then Theorem 3 observes the sequence of expected zero currents corresponding to the sequence of families  $\{P_N\}_{N \in \mathbb{N}}$  has a limit when normalized by a factor of  $N$ .

These observations conclude Section 2.1. As prefaced, we take an essential first step in our proof in Section 2.2.

## 2.2 Rewriting the function $F_n$

In Section 4.1 of “Random complex fewnomials, I,” an alternate expression for  $F_n$  is obtained and we duplicate the work here for both the sake of clarity and the subsequent use of the functions involved [SZ11]. Let  $\Delta$  be the unit simplex in  $\mathbb{R}^m$ .

## CHAPTER 2. BACKGROUND

For  $\rho \in \mathbb{R}^m$ , we write  $e^\rho = (e^{\rho_1}, \dots, e^{\rho_m})$  and  $|e^\rho| = \sum_{j=1}^m e^{\rho_j}$ . Define

$$b : \Delta \times \mathbb{R}^m \rightarrow [0, \infty) \quad (2.2.1)$$

$$(\lambda, \rho) \mapsto \langle \hat{\lambda}, \log \hat{\lambda} \rangle - \langle \rho, \lambda \rangle + \log(1 + |e^\rho|)$$

so that

$$F_n(\rho) = \log(1 + |e^\rho|) - \int_{\Delta^n} \min_{j=1, \dots, n} b(\lambda^j, \rho) d\lambda^1 \cdots d\lambda^n.$$

We proceed to use the following fact: Let  $X$  be a non-negative random variable on a probability space  $(\Omega, \mathcal{A}, \mathcal{P})$ , and let  $D_X(t) = \mathcal{P}(X \leq t)$  be its distribution function. Then the expected value of  $X$  is given by

$$\mathbb{E}(X) = \int X d\mathcal{P} = \int_0^\infty [1 - D_x(t)] dt. \quad (2.2.2)$$

If we let

$$D(t, \rho) = \mathcal{P}\{\lambda \in \Delta \mid b(\lambda, \rho) \leq t\}$$

be the distribution function for  $b(\lambda) = b(\lambda, \rho)$ , where  $d\mathcal{P}(\lambda) = d\lambda = m! d\lambda_1 \cdots d\lambda_m$ , then the distribution function for the random variable

$$X(\lambda^1, \dots, \lambda^n) = \min\{b(\lambda^1), \dots, b(\lambda^n)\} \quad (2.2.3)$$

on  $\Delta^n$ , with product measure  $d\mathcal{P}(\lambda^1) \cdots d\mathcal{P}(\lambda^n)$ , is given by

$$D_X = 1 - (1 - D)^n. \quad (2.2.4)$$

## CHAPTER 2. BACKGROUND

Combining (2.2.2) with (2.2.3) and (2.2.4), we obtain

$$F_n(\rho) - \log(1 + |e^\rho|) = - \int_{\Delta^n} X d\lambda^1 \cdots d\lambda^n = - \int_0^\infty [1 - D(t, \rho)]^n dt. \quad (2.2.5)$$

Having rewritten the potential function  $F_n$  as an integral depending on the distribution function  $D$  of  $b$ , we are one step closer to showing it is twice continuously differentiable. However, the function  $D$  is not useful to us in its current form and we are motivated to determine another way to write it.

The main idea for rewriting  $D$  can be found in Section 3.2 of Chapter 3, while the form of  $D$  we actually work with can be found in Section 4.1 of Chapter 4 where we begin putting all the pieces of Chapter 3 together.



## Chapter 3

# Notation, Functions and their Properties, and Motivation

In Chapter [3](#), we introduce notation, analyze the function  $b$  which occurs in the definition of  $D$ , discuss the case for  $m = 1$ , and generalize the discussion to higher dimension by introducing a function  $h$  which behaves like an inverse of  $b$ . We conclude the chapter by investigating the continuity of various partial derivatives for two of the functions which we introduce in Section [3.3](#) (namely the functions  $B$  and  $h$ ).

### 3.1 Notation and the function $b$

Let  $\Delta$  be the unit simplex in  $\mathbb{R}^m$ ; explicitly, we have

$$\begin{aligned}\Delta &= \{\lambda \in \mathbb{R}^m \mid \lambda_i \geq 0 \text{ for all } i = 1, \dots, m \text{ and } |\lambda| \leq 1\} \\ &= \{\lambda \in \mathbb{R}^m \mid \lambda_i \geq 0 \text{ for all } i = 0, 1, \dots, m\}\end{aligned}$$

where

$$|\lambda| \doteq \lambda_1 + \dots + \lambda_m$$

and

$$\lambda_0 \doteq 1 - |\lambda| \iff \sum_{i=0}^m \lambda_i = 1. \quad (3.1.1)$$

Then let  $\Delta^\circ$  denote its interior and  $\partial\Delta$  denote its boundary. For  $i = 0, 1, \dots, m$ , define the  $i$ -th facet of  $\Delta$  to be  $\partial^i\Delta \doteq \{\lambda \in \Delta \mid \lambda_i = 0\}$ . It follows that  $\partial\Delta$  is a union of the facets  $\partial^i\Delta$ . For  $i \in \{0, 1, \dots, m\}$ , we also define  $v_i \in \Delta$  to be the vertex of  $\Delta$  which is nonadjacent to the  $i$ -th facet. By definition, the  $i$ -th vertex  $v_i$  is the point  $\lambda \in \Delta$  with  $\lambda_i = 1$  (equivalently, we may write  $v_i = (\delta_{1i}, \delta_{2i}, \dots, \delta_{mi})$  where  $\delta_{ij}$  is the Kronecker delta function).

Now write  $e^\rho = (e^{\rho_1}, \dots, e^{\rho_m})$  and define the (surjective) map

$$\begin{aligned}\mu : \mathbb{R}^m &\rightarrow \Delta^\circ \\ \rho &\mapsto \frac{e^\rho}{1 + |e^\rho|}.\end{aligned}$$

The map  $\mu$  is the moment map of  $\mathbb{CP}^m$  with respect to the Fubini-Study metric [MR02,

### CHAPTER 3. NOTATION, FUNCTIONS, AND MOTIVATION

[SZ04](#)]. Denoting the  $i$ -th component of  $\mu$  by  $\mu_i$ , we have  $\mu(\rho) = (\mu_1(\rho), \dots, \mu_m(\rho))$

and by (3.1.1) we obtain

$$\mu_0(\rho) = 1 - |\mu(\rho)| = \frac{1}{1 + |e^\rho|}.$$

Defining  $\rho_0 = 0$ , then we have

$$\mu_i(\rho) = \frac{e^{\rho_i}}{1 + |e^\rho|}$$

for  $i = 0, 1, \dots, m$ . We remark that the ability to generalize the formula above to include the case for  $i = 0$  comes from associating a polynomial  $\sum_{|\alpha| \leq N} c_\alpha z^\alpha$  on  $\mathbb{C}^m$  with the homogeneous polynomial  $\sum_{|\alpha| \leq N} c_\alpha z_0^{\alpha_0} z^\alpha$  on  $\mathbb{C}^{m+1}$ . Then the original polynomial is obtained by restricting the homogeneous polynomial to  $z_0 = 1$ . Since  $\rho_i = \log |z_i|^2$ , we have  $\rho_0 = 0$ .

For  $i = 1, \dots, m$ , we compute the first partial derivatives of  $\mu_i = \mu_i(\rho)$  with respect to  $\rho_p$ . When  $p = i$ , we have

$$\mu_{i,i} \doteq \frac{\partial}{\partial \rho_i} \mu_i = \frac{e^{\rho_i}}{1 + |e^\rho|} - \frac{e^{\rho_i} e^{\rho_i}}{(1 + |e^\rho|)^2}.$$

When  $p \neq i$ , we have

$$\mu_{i,p} \doteq \frac{\partial}{\partial \rho_p} \mu_i = -\frac{e^{\rho_i} e^{\rho_p}}{(1 + |e^\rho|)^2}.$$

And last, but not least, for  $i = 0$ , we have

$$\mu_{0,p} \doteq \frac{\partial}{\partial \rho_p} \mu_0 = -\frac{e^{\rho_0} e^{\rho_p}}{(1 + |e^\rho|)^2}.$$

Conveniently, we summarize the above computations into one succinct formula and

### CHAPTER 3. NOTATION, FUNCTIONS, AND MOTIVATION

write

$$\mu_{i,p} = \mu_{i\&p} - \mu_i \mu_p, \quad (3.1.2)$$

where

$$\begin{aligned} \mu_{i\&j} &\doteq \begin{cases} \mu_i & \text{if } i = j, \\ 0 & \text{otherwise.} \end{cases} \\ &= \delta_{ij} \mu_i \end{aligned}$$

As a side note, we observe that the indices exhibit a symmetry (e.g.,  $\mu_{i,p} = \mu_{p,i}$  for  $i \neq 0$ ) because  $\mu_i = \frac{\partial}{\partial \rho_i} \log(1 + |e^\rho|)$ .

Next we analyze the function  $b$  which we defined at (2.2.1).

**Lemma 5.** *The function*

$$\begin{aligned} b(\lambda, \rho) &= \langle \hat{\lambda}, \log \hat{\lambda} \rangle - \langle \rho, \lambda \rangle + \log(1 + |e^\rho|) \\ &= \log(1 + |e^\rho|) + \sum_{i=0}^m [\lambda_i \log \lambda_i - \rho_i \lambda_i], \end{aligned}$$

for  $(\lambda, \rho) \in \Delta \times \mathbb{R}^m$ , has the following properties:

- 1)  $b$  is smooth on  $\Delta^\circ \times \mathbb{R}^m$ ,
- 2)  $b$  is strictly convex with respect to the variable  $\lambda$ ,
- 3)  $b(\lambda, \rho) = 0$  if and only  $\lambda = \mu(\rho)$ , and
- 4)  $\max_{\lambda \in \Delta_m} b(\lambda, \rho) = \log(1 + |e^\rho|) - \min\{\rho_0, \rho_1, \dots, \rho_m\}$ .

*Proof.* The first property is clear.

### CHAPTER 3. NOTATION, FUNCTIONS, AND MOTIVATION

To show that  $b$  is strictly convex with respect to  $\lambda$ , we compute the Hessian of  $b(\cdot, \rho)$ . Alternatively, one can also check the definition of convexity and make use of the fact that  $x \mapsto x \log x$  is a convex function for  $x > 0$ . In any case, we move forward and compute the first partial derivatives of  $b$  in  $\lambda$  to be

$$b_{\lambda_i}(\lambda, \rho) = \log \lambda_i - \log \lambda_0 - \rho_i, \quad \text{for } (\lambda, \rho) \in \Delta^\circ \times \mathbb{R}^m$$

and we compute the second partial derivatives of  $b$  in  $\lambda$  to be

$$b_{\lambda_i \lambda_j}(\lambda, \rho) = \frac{1}{\lambda_i} \delta_{ij} + \frac{1}{\lambda_0}, \quad \text{for } (\lambda, \rho) \in \Delta^\circ \times \mathbb{R}^m.$$

Thus, the Hessian of  $b(\cdot, \rho)$  is given by

$$H(\lambda) = \begin{bmatrix} \lambda_1^{-1} & 0 & \cdots & 0 \\ 0 & \lambda_2^{-1} & \cdots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & 0 & \lambda_m^{-1} \end{bmatrix} + \lambda_0^{-1} \begin{bmatrix} 1 & 1 & \cdots & 1 \\ 1 & 1 & \cdots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & \cdots & 1 \end{bmatrix}.$$

Then for any  $v \in \mathbb{R}^m \setminus \{0\}$  and  $\lambda \in \Delta^\circ$ , we have

$$v^\top H(\lambda) v = \sum_{i=1}^m \lambda_i^{-1} v_i^2 + \lambda_0^{-1} \left[ \sum_{i=1}^m v_i \right]^2 > 0$$

and conclude the Hessian of  $b(\cdot, \rho)$  is positive definite on the interior of the unit simplex. Hence,  $b(\cdot, \rho)$  is strictly convex on the unit simplex.

For the third property, it is an immediate check that  $b(\mu(\rho), \rho) = 0$  and we are left to show the forward implication. We fix  $\rho \in \mathbb{R}^m$  and solve the system

## CHAPTER 3. NOTATION, FUNCTIONS, AND MOTIVATION

of equations  $b_{\lambda_i}(\lambda, \rho) = 0$  in order to determine critical points (for maxima and minima) of  $b(\cdot, \rho)$ . For  $i = 1, \dots, m$ , we immediately obtain  $\lambda_i = \lambda_0 e^{\rho_i}$ . Together with  $\sum_{i=0}^m \lambda_i = 1$ , we have  $\lambda_0 + \sum_{i=1}^m (\lambda_0 e^{\rho_i}) = 1$  and we obtain  $\lambda_0 = \mu_0(\rho)$ . Thus, we have  $\lambda_i = \mu_0(\rho) e^{\rho_i} = \mu_i(\rho)$  for  $i = 1, \dots, m$  and we conclude  $b(\cdot, \rho)$  has a critical point at  $\mu(\rho)$ . We deduce from the strict convexity of  $b(\cdot, \rho)$  that  $b(\cdot, \rho)$  has a unique global minimum at  $\mu(\rho)$ . It follows that  $b(\lambda, \rho) = 0$  only if  $\lambda = \mu(\rho)$ , as desired.

Finally, because  $b$  is convex on  $\Delta$ , it must take its maximum at one of the vertices  $v_i$  of  $\Delta$ . For  $i = 0, 1, \dots, m$ , we have  $b(v_i, \rho) = \log(1 + |e^\rho|) - \rho_i$ . Therefore, we take the maximum over  $i = 0, 1, \dots, m$  and obtain the fourth and final property.  $\square$

## 3.2 Motivation

While the cases for  $m = 1$  and  $m = 2$  do not require special attention, our ability to visualize real-valued functions on  $\mathbb{R}$  and  $\mathbb{R}^2$  come in handy to understand the work done for arbitrary dimensions.

For  $m = 1$ , Lemma 5 implies the function

$$b(\lambda, \rho) = \lambda \log \lambda + (1 - \lambda) \log(1 - \lambda) - \lambda \rho + \log(1 + e^\rho)$$

is convex in  $\lambda$  with an absolute minimum at  $\lambda = \mu(\rho) = e^\rho[1 + e^\rho]^{-1}$ . Consequently, the function  $b(\cdot, \rho)$  has an inverse branch  $g_{-1}(\cdot, \rho)$  on the interval  $[0, \mu(\rho)]$  and an inverse branch  $g_1(\cdot, \rho)$  on the interval  $[\mu(\rho), 1]$ .

Referring to Figure 3.2.1, the graph for  $b(\lambda, \rho)$  is plotted on the left for  $\rho = 0.5$

### CHAPTER 3. NOTATION, FUNCTIONS, AND MOTIVATION

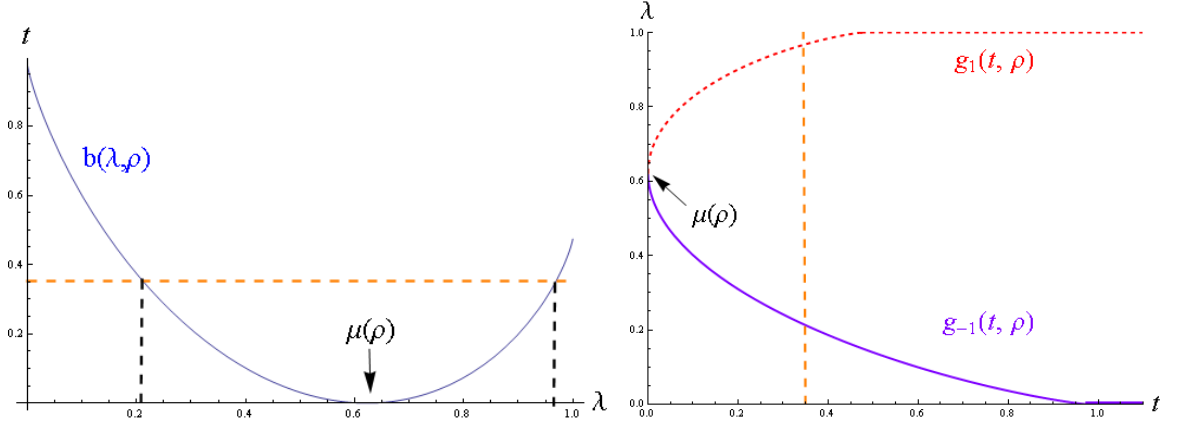


Figure 3.2.1: The functions  $b(\lambda, \rho)$  and its two inverse branches  $g_1(t, \rho)$  and  $g_{-1}(t, \rho)$  for  $\rho = 0.5$ .

and its inverse branches are plotted on the right. Determining the value of  $D(t_0, \rho)$  graphically is then accomplished by drawing a dashed line at  $t = t_0$ , determining where the line intersects with the inverse branches, and then taking the difference. For example, when  $t = 0.35$  and  $\rho = 0.5$ , then the dashed line at  $t = 0.35$  in Figure 3.2.1 crosses the inverse branch  $g_{-1}$  at approximately  $\lambda = 0.20$  and crosses the inverse branch  $g_1$  at approximately  $\lambda = 0.96$ . Thus  $D(0.35, 0.5)$  is approximately  $0.96 - 0.20 = 0.76$ . Actually computing  $D(0.35, 0.5)$  gives the approximate value of .755712.

Keeping the above example in mind, we take the branches  $g_1$  and  $g_{-1}$  and extending their domain continuously to  $[0, \infty)$ . From this we obtain

$$g_{-1}(t, \rho) = \begin{cases} \lambda, & \text{if } t \in [0, b(0, \rho)], b(\lambda, \rho) = t, \text{ and } \lambda \in [0, \mu(\rho)] \\ 0, & \text{if } t \in [b(0, \rho), \infty) \end{cases}$$

and

### CHAPTER 3. NOTATION, FUNCTIONS, AND MOTIVATION

$$g_1(t, \rho) = \begin{cases} \lambda, & \text{if } t \in [0, b(1, \rho)], b(\lambda, \rho) = t, \text{ and } \lambda \in [\mu(\rho), 1] \\ 1, & \text{if } t \in [b(1, \rho), \infty), \end{cases}$$

so that

$$\begin{aligned} D(t, \rho) &= \mathcal{P}\{\lambda \in \Delta \mid b(\lambda, \rho) \leq t\} \\ &= g_1(t, \rho) - g_{-1}(t, \rho). \end{aligned} \tag{3.2.1}$$

We note that the above evaluation of  $D$  as a difference of inverse branches appears in Section 4.1.1 of “Random complex fewnomials, I” [SZ11], but the article does not generalize the decomposition to higher dimensions.

When looking to generalize the idea to higher dimensions, we further write

$$\begin{aligned} D(t, \rho) &= g_1(t, \rho) - \mu(\rho) + \mu(\rho) - g_{-1}(t, \rho) \\ &= |g_1(t, \rho) - \mu(\rho)| + |g_{-1}(t, \rho) - \mu(\rho)|. \end{aligned}$$

With the above in mind, we begin consider the case for  $m = 2$ . However, before going into details, we look at Figure 3.2.2 which shows a graph (left) and 10-level contour plot (right) for the function  $b(\cdot, \rho_0)$  where  $\rho_0 = (0.5, 0)$ . Letting  $M$  being the maximum value attained by  $b(\cdot, \rho_0)$ , then the value of  $D(t_0, \rho_0)$  for  $t_0 = M/10$  is precisely given by the area bounded by the first contour about  $\mu(\rho_0)$ . Likewise, the value of  $D(k \cdot t_0, \rho_0)$  is given by the area bounded by the  $k$ th contour away from the point  $\mu(\rho_0)$ .



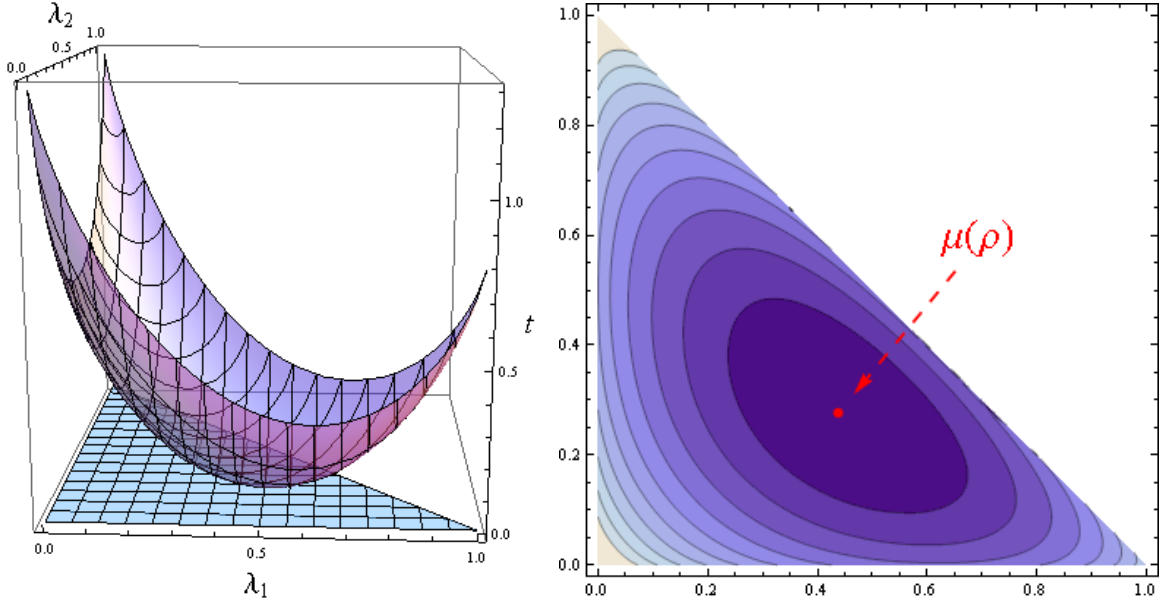


Figure 3.2.2: A 3D plot and a 10-level contour plot of  $b(\lambda, \rho)$  for  $\rho = (0.5, 0)$ .

By the convexity of  $b(\cdot, \rho_0)$  and imposing a system of polar coordinates centered at  $\mu(\rho_0)$ , the first contour, given as the set of points  $\lambda \in \mathbb{R}^2$  such that  $b(\lambda, \rho_0) = t_0$ , can be parametrized by a polar equation  $\theta \mapsto r(\theta)$ , see Figure 3.2.3. More explicitly, given  $\theta \in \mathbb{R}$ , if the ray with vertex  $\mu(\rho_0)$  which forms an angle  $\theta$  counter-clockwise from the horizontal intersects the contour at a point  $a(\theta)$ , then  $r(\theta) = |a(\theta) - \mu(\rho_0)|$ . Recalling some calculus, the area bounded by the first contour is given by  $\frac{1}{2} \int_0^{2\pi} r(\theta)^2 d\theta$ .

However, continuing to look at Figure 3.2.3, we can make a change of variables so that we parametrize over the boundary of the unit simplex instead of parametrizing over the angles  $\theta$ .

Rigorously, Lemma 5 implies the function  $b(\cdot, \rho)$  is convex with an absolute minimum at  $\lambda = \mu(\rho)$ . Consequently, given any unit direction  $u \in \mathbb{R}^2$ , the function  $b(\cdot, \rho)$  has an inverse branch  $g_u(\cdot, \rho)$  with respect to the direction  $u$ . Taking these branches

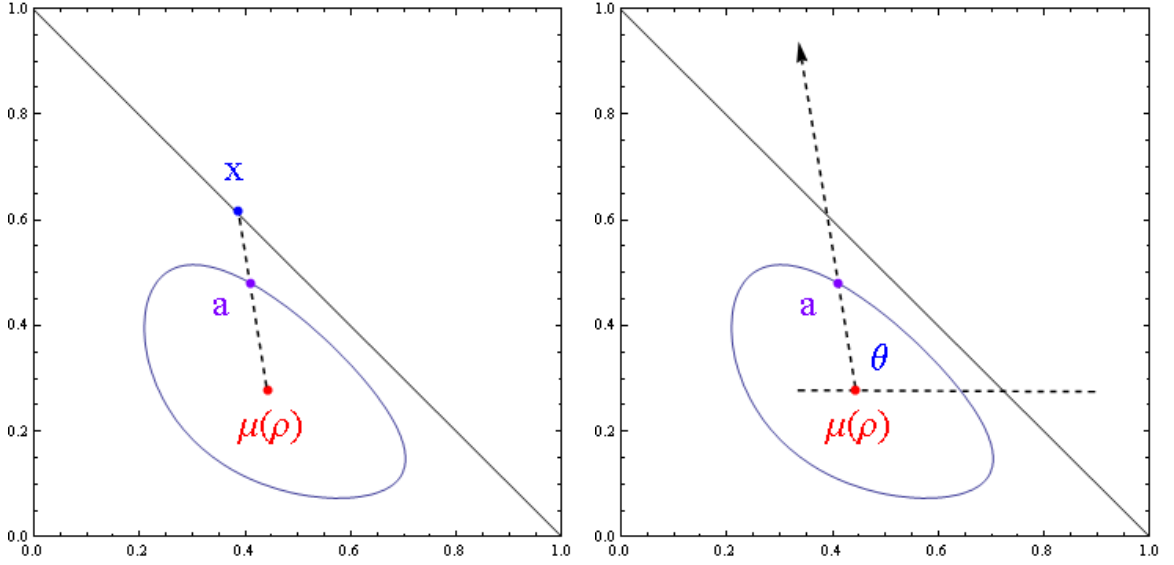


Figure 3.2.3: The ray with vertex  $\mu(\rho_0)$  and angle  $\theta$  intersects the curve  $b(\lambda, \rho_0) = t_0$  at the point  $a$ . Equivalently, there is a point  $x$  on the boundary such that the line segment from  $\mu(\rho)$  to a point  $x$  also intersects the curve at  $a$ .

and extending their domain continuously to  $[0, \infty)$ , we obtain

$$g_u(t, \rho) = \begin{cases} \lambda, & \text{if } t \in [0, b(x_u, \rho)], \ b(\lambda, \rho) = t, \text{ and } \lambda = \mu + su \text{ for some } s \geq 0 \\ x_u & \text{if } t \in [b(x_u, \rho), \infty), \end{cases}$$

where  $x_u$  is the unique element in the boundary of the unit simplex  $\partial\Delta$  such that  $x_u = \mu + s_u u$  for some  $s_u > 0$ . (We note that such an  $x_u$  exists and is unique because the unit simplex  $\Delta$  is convex and  $\mu \in \Delta$ .) Then, loosely speaking, the function  $D$  in the case for  $m = 2$  would be the integral of the lengths  $\|g_u(t, \rho) - \mu(\rho)\|$  over all unit directions  $u$  in  $\mathbb{R}^2$ . However, the definition of  $g_u$  is quite cumbersome and we are actually interested in the lengths  $\|g_u(t, \rho) - \mu(\rho)\|$ . Thus, a small improvement is to

### CHAPTER 3. NOTATION, FUNCTIONS, AND MOTIVATION

consider the functions

$$\tilde{f}_u(t, \rho) = \begin{cases} s, & \text{if } t \in [0, b(x_u, \rho)] \text{ and } b(\mu + su, \rho) = t \text{ for some } s \geq 0 \\ s_u, & \text{if } t \in [b(x_u, \rho), \infty), \end{cases} \quad (3.2.2)$$

where  $x_u$  is the unique element in the boundary of the unit simplex  $\partial\Delta$  such that  $x_u = \mu + s_u u$  for some  $s_u > 0$ . This will be our initial approach in Section 3.3 as it allows us to apply the implicit function theorem. For  $m = 1$ , we have

$$D(t, \rho) = \tilde{f}_1(t, \rho) + \tilde{f}_{-1}(t, \rho)$$

However, when we want to actually take derivatives and set up an integral, it is convenient to parametrize using the boundary points  $x$  in  $\partial\Delta$ , instead of using unit directions  $u$  in  $\mathbb{R}^2$ , see Figure 3.2.3. Explicitly, we consider the functions

$$f_x(t, \rho) = \begin{cases} s, & \text{if } t \in [0, b(x, \rho)] \text{ and } b((1-s)\mu + sx, \rho) = t \\ 1, & \text{if } t \in [b(x, \rho), \infty). \end{cases} \quad (3.2.3)$$

Note, however, the formulation of this family of functions  $\{f_x\}_{x \in \partial\Delta}$  is simpler than the formulation of the family of functions  $\{\tilde{f}_u\}_{u \in S^m}$ . In addition, the range of  $f_x(\cdot, \rho)$ , for any  $x \in \partial\Delta$ , is the interval  $[0, 1]$ , while the range of  $\tilde{f}_u(\cdot, \rho)$  varies as  $u$  varies. Noting that  $(1-s)\mu + sx = \mu + s(x - \mu)$ , we obtain

$$f_x(t, \rho) = \frac{1}{\|x - \mu\|} \tilde{f}_u(t, \rho), \quad \text{where } u = \frac{x - \mu}{\|x - \mu\|}, \quad (3.2.4)$$

## CHAPTER 3. NOTATION, FUNCTIONS, AND MOTIVATION

or equivalently

$$\tilde{f}_u(t, \rho) = \|x_u - \mu\| f_{x_u}(t, \rho), \quad \text{where } x_u = \mu + su \in \partial\Delta, \ s > 0,$$

which gives a one-to-one correspondence between the family of functions  $\{\tilde{f}_u(\cdot, \rho)\}_{u \in S^1}$  and the family of functions  $\{f_x(\cdot, \rho)\}_{x \in \partial\Delta}$ . For  $m = 1$ , we have

$$\begin{aligned} D(t, \rho) &= \tilde{f}_1(t, \rho) + \tilde{f}_{-1}(t, \rho) \\ &= |1 - \mu_1| f_1(t, \rho) + |0 - \mu_1| f_0(t, \rho) \\ &= \mu_0 \cdot f_1(t, \rho) + \mu_1 \cdot f_0(t, \rho). \end{aligned} \tag{3.2.5}$$

This completes our motivation for the functions we plan to define.

### 3.3 Defining the functions $\Lambda$ , $B$ , and $h$

First, we define some auxiliary functions, which after this section, we no longer consider. Specifically, we will set out to rigorously define a function  $\tilde{h}$ , in fact  $\tilde{h}(t, \rho, v) = \tilde{f}_{v/\|v\|}(t, \rho)$  (3.2.2), only to rigorously define a function  $h$ , in fact  $h(t, \rho, x) = f_x(t, \rho)$  (3.2.3), which inherits desired smoothness properties from  $h$  by (3.2.4).

We start by defining the (surjective) map

$$\begin{aligned} \tilde{\Lambda} : \mathbb{R} \times \mathbb{R}^m \times (\mathbb{R}^m \setminus \{0\}) &\rightarrow \mathbb{R}^m \\ (\tilde{s}, \rho, v) &\mapsto \mu(\rho) + \tilde{s} \frac{v}{\|v\|} \end{aligned}$$

### CHAPTER 3. NOTATION, FUNCTIONS, AND MOTIVATION

and consider the function

$$\tilde{B}(\tilde{s}, \rho, v) = b(\tilde{\Lambda}(\tilde{s}, \rho, v), \rho),$$

whose domain we denote by  $\tilde{S}$ . By Lemma 5, the function  $\tilde{B}$  is smooth on the interior of  $\tilde{S}$  and  $\tilde{B}(\cdot, \rho, v)$  is a strictly increasing function for  $\tilde{s} > 0$ . Combining these facts with the implicit function theorem, there exists an open set  $\tilde{H}$  in  $(0, \infty) \times \mathbb{R}^m \times (\mathbb{R}^m \setminus \{0\})$  and a smooth function  $\tilde{h} : \tilde{H} \rightarrow [0, \infty)$  such that

$$\tilde{B}(\tilde{h}(t, \rho, v), \rho, v) = t, \quad \text{for } (t, \rho, v) \in \tilde{H}. \quad (3.3.1)$$

For the sake of completeness, we give a description of the set  $\tilde{H}$ , but note that it is not essential, but note that it has little importance. For  $\rho \in \mathbb{R}^m$ ,  $v \in \mathbb{R}^m \setminus \{0\}$ , and  $(t, \rho, v) \in \tilde{H}$ , then  $t$  must belong to the interior of the range of  $\tilde{B}(\cdot, \rho, v)|_{\tilde{s} \geq 0}$ . In other words, we have

$$\tilde{H} = \{(t, \rho, v) \mid \tilde{B}(\tilde{s}, \rho, v) = t \text{ for some } (\tilde{s}, \rho, v) \in \tilde{S} \text{ with } \tilde{s} > 0\}.$$

Now fix  $x \in \partial\Delta$ , take (3.3.1), and set  $v = x - \mu(\rho) \in \mathbb{R}^m \setminus \{0\}$ . Then we have

$$b\left(\mu + \tilde{h}(t, \rho, x - \mu) \frac{x - \mu}{\|x - \mu\|}, \rho\right) = t,$$

or

$$b(\mu + h(t, \rho, x)(x - \mu), \rho) = t, \quad \text{where } h(t, \rho, x) = \frac{1}{\|x - \mu\|} \tilde{h}(t, \rho, x - \mu).$$

Similar to replacing  $\{f_u\}_{u \in S^1}$  with  $\{f_x\}_{x \in \Delta}$  in Section 3.2, we will replace  $\tilde{\Lambda}$ ,  $\tilde{B}$ ,

### CHAPTER 3. NOTATION, FUNCTIONS, AND MOTIVATION

$\tilde{H}$ , and  $\tilde{h}$ , with their analogs  $\Lambda$ ,  $B$ ,  $H$ , and  $h$ . We define the (surjective) map

$$\begin{aligned}\Lambda &: [0, 1] \times \mathbb{R}^m \times \partial\Delta \twoheadrightarrow \Delta \\ (s, \rho, x) &\mapsto \mu(\rho) + s(x - \mu(\rho)) \\ &= (1 - s)\mu(\rho) + sx \quad ,\end{aligned}$$

the function

$$\begin{aligned}B &: [0, 1] \times \mathbb{R}^m \times \partial\Delta \rightarrow [0, \infty) \\ (s, \rho, x) &\mapsto b(\Lambda(s, \rho, x), \rho),\end{aligned}$$

the domain

$$H = \{(t, \rho, x) \in (0, \infty) \times \mathbb{R}^m \times \partial\Delta \mid 0 < t < b(x, \rho)\},$$

and the function

$$\begin{aligned}h &: H \twoheadrightarrow (0, 1) \\ (t, \rho, x) &\mapsto s \text{ such that } B(s, \rho, x) = t.\end{aligned}$$

Using Lemma 5 and the observation that the mixed partial derivatives of the map  $\Lambda$  in  $s$  and  $\rho$  of all orders are continuous on  $[0, 1] \times \mathbb{R}^m \times \partial\Delta$ , we find the mixed partial derivatives of  $B$  in  $s$  and  $\rho$  (of all orders) are continuous on  $[0, 1] \times \mathbb{R}^m \times \partial\Delta$ . Lemma 5 also implies that  $B$  is a strictly increasing function with respect to  $s$ .

Similarly for  $h$ , we note that the partial derivatives of the maps  $(\rho, x) \mapsto x - \mu(\rho)$

## CHAPTER 3. NOTATION, FUNCTIONS, AND MOTIVATION

and  $(\rho, x) \mapsto \|x - \mu(\rho)\|^{-1}$  in  $\rho$  of all orders are continuous on  $\mathbb{R}^m \times \partial\Delta$ . Then the mixed partial derivatives of  $h$  in  $t$  and  $\rho$  (of all orders) are continuous on  $H$ , since  $\tilde{h} \in \mathcal{C}^\infty(\tilde{H})$ . We also have the function  $h$  is strictly increasing with respect to  $t$ , because the function  $B$  is strictly increasing with respect to  $s$ .

### 3.4 Properties of the function $B$

As the definition of  $h$  is implicitly defined by  $B$ , we proceed by first investigating  $B$  and its derivatives along the boundary of its domain. A summary may be found in Section 3.4.3.

#### 3.4.1 Along $s = 0$ .

As mentioned in Section 3.3, the mixed partial derivatives of  $B$  in  $s$  and  $\rho$  (of all orders) are continuous on  $[0, 1) \times \mathbb{R}^m \times \partial\Delta$ . Thus, along  $s = 0$ , we simply look to evaluate the derivatives of  $B$ .

The first useful fact in our computations follows from the definition of  $\mu_0$  and  $\mu_i$ . For  $i = 0, 1, \dots, m$ , we have

$$\log[\mu_0] = \log[\mu_i] - \rho_i. \quad (3.4.1)$$

Taking the first partial derivative of  $B$  with respect to  $s$ , we obtain

$$B_s(s, \rho, x) = \sum_{i=0}^m \left[ \log[(1-s)\mu_i + sx_i] - \rho_i \right] (x_i - \mu_i)$$

### CHAPTER 3. NOTATION, FUNCTIONS, AND MOTIVATION

and note that it is non-negative, equaling 0 if and only if  $s = 0$ , because  $B$  is strictly increasing for  $s > 0$ . Checking  $B_s(0, \rho, x) = 0$ , we write

$$\begin{aligned} B_s(0, \rho, x) &= \sum_{i=0}^m \left[ \log[\mu_i] - \rho_i \right] (x_i - \mu_i) \\ &\stackrel{3.4.1}{=} \log[\mu_0] \cdot \sum_{i=0}^m (x_i - \mu_i) = 0, \end{aligned} \quad (3.4.2)$$

where the last equality is given by

$$\sum_{i=0}^m (x_i - \mu_i) = \sum_{i=0}^m x_i - \sum_{i=0}^m \mu_i \stackrel{3.1.1}{=} 1 - 1 = 0. \quad (3.4.3)$$

While evaluating  $B_s$  at  $s = 0$  took several steps, no heavy computation is required for evaluating  $B_{ss}$  along  $s = 0$ . We have

$$B_{ss}(s, \rho, x) = \sum_{i=0}^m \frac{(x_i - \mu_i)^2}{(1-s)\mu_i + sx_i} > 0,$$

and so

$$B_{ss}(0, \rho, x) = \sum_{i=0}^m \frac{(x_i - \mu_i)^2}{\mu_i}. \quad (3.4.4)$$

Now let us differentiate  $B$  with respect to  $\rho_p$ . As we apply the chain rule, we note that

$$b_{\rho_i}(\lambda, \rho) = -\lambda_i + \mu_i(\rho), \quad \text{for } (\lambda, \rho) \in \Delta \times \mathbb{R}^m$$

and so

$$B_{\rho_p}(s, \rho, x) = -s(x_p - \mu_p) + (1-s) \sum_{i=0}^m \left[ \log[(1-s)\mu_i + sx_i] - \rho_i \right] \mu_{i,p}.$$



### CHAPTER 3. NOTATION, FUNCTIONS, AND MOTIVATION

Evaluating the derivative for  $s = 0$ , we have

$$\begin{aligned} B_{\rho_p}(0, \rho, x) &= \sum_{i=0}^m \left[ \log[\mu_i] - \rho_i \right] \mu_{i,p} \\ &\stackrel{3.4.1}{=} \log[\mu_0] \cdot \sum_{i=0}^m \mu_{i,p} = 0, \end{aligned}$$

where the last equality is given by

$$\sum_{i=0}^m \mu_{i,p} = \frac{\partial}{\partial \rho_p} \left[ \sum_{i=0}^n \mu_i \right] \stackrel{3.1.1}{=} 0. \quad (3.4.5)$$

Next, we consider the partial derivative

$$\begin{aligned} B_{\rho_p s}(s, \rho, x) &= -(x_p - \mu_p) \\ &\quad + \sum_{i=0}^m \left[ -\log[(1-s)\mu_i + sx_i] + \rho_i + \frac{(1-s)(x_i - \mu_i)}{(1-s)\mu_i + sx_i} \right] \mu_{i,p} \end{aligned}$$

so that

$$\begin{aligned} B_{\rho_p s}(0, \rho, x) &= -(x_p - \mu_p) + \sum_{i=0}^m \left[ -\log[\mu_i] + \rho_i + \frac{(x_i - \mu_i)}{\mu_i} \right] \mu_{i,p} \\ &\stackrel{3.4.1}{=} -(x_p - \mu_p) + \sum_{i=0}^m \left[ -\log[\mu_0] + \frac{(x_i - \mu_i)}{\mu_i} \right] \mu_{i,p} \\ &\stackrel{3.4.5}{=} -(x_p - \mu_p) + \sum_{i=0}^m \left[ \frac{(x_i - \mu_i)}{\mu_i} \right] \mu_{i,p} = 0 \end{aligned} \quad (3.4.6)$$

where the last equality is given by

$$\sum_{i=0}^m \frac{x_i - \mu_i}{\mu_i} \mu_{i,p} \stackrel{3.1.2}{=} \sum_{i=0}^m (x_i - \mu_i)(\delta_{ip} - \mu_p) \stackrel{3.4.3}{=} x_p - \mu_p.$$

## CHAPTER 3. NOTATION, FUNCTIONS, AND MOTIVATION

Finally, we have the partial derivative

$$B_{\rho_p \rho_q}(s, \rho, x) = (2s - 1)\mu_{p,q} + \sum_{i=0}^m \frac{\mu_{i,q}\mu_{i,p} \cdot (1-s)^2}{(1-s)\mu_i + sx_i} \\ + (1-s) \sum_{i=0}^m \left[ \log[(1-s)\mu_i + sx_i] - \rho_i \right] \mu_{i,p,q}$$

for which

$$B_{\rho_p \rho_q}(0, \rho, x) = -\mu_{p,q} + \sum_{i=0}^m \frac{\mu_{i,q}\mu_{i,p}}{\mu_i} + \sum_{i=0}^m \left[ \log[\mu_i] - \rho_i \right] \mu_{i,p,q} \\ \stackrel{3.4.1}{=} -\mu_{p,q} + \sum_{i=0}^m \frac{\mu_{i,q}\mu_{i,p}}{\mu_i} + \log[\mu_0] \cdot \sum_{i=0}^m \mu_{i,p,q} \\ = -\mu_{p,q} + \sum_{i=0}^m \mu_{i,q} \cdot \frac{\mu_{i,p}}{\mu_i} = 0,$$

where the third equality is given by

$$\sum_{i=0}^m \mu_{i,p,q} = \frac{\partial^2}{\partial \rho_p \partial \rho_q} \left[ \sum_{i=0}^m \mu_i \right] \stackrel{3.1.1}{=} 0,$$

and the last equality is given by

$$\sum_{i=0}^m \mu_{i,q} \cdot \frac{\mu_{i,p}}{\mu_i} \stackrel{3.1.2}{=} \sum_{i=0}^m \mu_{i,q} \cdot (\delta_{ip} - \mu_p) \stackrel{3.4.5}{=} \mu_{p,q}.$$

### 3.4.2 Along $s = 1$ .

In contrast to Section 3.4.1 where we simply evaluate various partial derivatives of  $B$  along  $s = 0$ , this section deals with the limiting behavior of the partial derivatives as  $s$  tends to 1.

Then let us begin by fixing  $\tilde{\rho} \in \mathbb{R}^m$  and  $\tilde{x} \in \partial\Delta$ . We also consider the variables

### CHAPTER 3. NOTATION, FUNCTIONS, AND MOTIVATION

$s \in (0, 1)$ ,  $\rho \in \mathbb{R}^m$ , and  $x \in \partial\Delta$ .

For the limit of  $B_s$  as  $(s, \rho, x)$  tends to  $(1, \tilde{\rho}, \tilde{x})$ , we note that the summands  $(x_i - \mu_i) \log[(1 - s)\mu_i + sx_i]$  of  $B_s$ , for all  $i = 0, 1, \dots, m$  such that  $\tilde{x}_i \neq 0$ , and the summands  $-\rho_i(x_i - \mu_i)$  of  $B_s$ , for all  $i = 0, 1, \dots, m$ , have finite limits as  $(s, \rho, x)$  tends to  $(1, \tilde{\rho}, \tilde{x})$ . We next consider the subset  $S \subset \{0, 1, \dots, m\}$  of indices for which  $\tilde{x}_i = 0$ . Then there exists  $\delta_1 > 0$  such that  $x_i - \mu_i < 0$  for all  $i \in S$  and  $\|(s, \rho, x) - (1, \tilde{\rho}, \tilde{x})\| < \delta_1$ . Thus all the summands of  $\sum_{i \in S} (x_i - \mu_i) \log[(1 - s)\mu_i + sx_i]$  are positive and we have

$$\begin{aligned} \sum_{i \in S} (x_i - \mu_i) \log[(1 - s)\mu_i + sx_i] &\geq (x_k - \mu_k) \log[(1 - s)\mu_k + sx_k] \\ &= -\mu_k \cdot \log[(1 - s)\mu_k], \end{aligned}$$

where  $k$  is an index for which  $x_k = 0$ . For any  $M > 0$ , there exists  $\delta_2$  such that  $-\mu_k \cdot \log[(1 - s)\mu_k] > M$  for all  $k \in S$  and  $\|(s, \rho, x) - (1, \tilde{\rho}, \tilde{x})\| < \delta_2$ . We conclude that

$$\lim_{(s, \rho, x) \rightarrow (1, \tilde{\rho}, \tilde{x})} B_s(s, \rho, x) = \infty. \quad (3.4.7)$$

For the limit of  $B_{ss}$  as  $(s, \rho, x)$  tends to  $(1, \tilde{\rho}, \tilde{x})$ , we first note that all the summands are positive and proceed similar to the proof of (3.4.7). We have

$$\begin{aligned} B_{ss}(s, \rho, y) &\geq (x_k - \mu_k)^2 [(1 - s)\mu_k + sx_k]^{-1} \\ &= (-\mu_k)^2 [(1 - s)\mu_k]^{-1} = (1 - s)^{-1} \mu_k, \end{aligned}$$

where  $k$  is an index for which  $x_k = 0$ . For any  $M > 0$ , there exists a  $\delta > 0$  such that

### CHAPTER 3. NOTATION, FUNCTIONS, AND MOTIVATION

$\mu_k \cdot (1-s)^{-1} > M$  for all  $k \in \{0, 1, \dots, m\}$  and  $\|(s, \rho, x) - (1, \rho^0, x^0)\| < \delta$ . Thus, we have

$$\lim_{(s, \rho, x) \rightarrow (1, \tilde{\rho}, \tilde{x})} B_{ss}(s, \rho, x) = \infty.$$

Next we look at the limit of  $B_{\rho_p}$  as  $(s, \rho, x)$  tends to  $(1, \tilde{\rho}, \tilde{x})$ . We note that the summation  $\sum_{i=0}^m (1-s)(-\rho_i)\mu_{i,\rho}$  tends to 0 and the term  $-s(x_p - \mu_p)$  tends to  $-\tilde{x}_p - \mu_p(\tilde{\rho})$  as  $(s, \rho, x)$  tends to  $(1, \tilde{\rho}, \tilde{x})$ . Thus it remains to investigate the limit of  $\sum_{i=0}^m (1-s)\mu_{i,p} \cdot \log[(1-s)\mu_i + sx_i]$  as  $(s, \rho, x)$  tends to  $(1, \tilde{\rho}, \tilde{x})$ . We have

$$\begin{aligned} \left| (1-s) \sum_{i=0}^m \mu_{i,p} \cdot \log[(1-s)\mu_i + sx_i] \right| &\leq (1-s) \sum_{i=0}^m |\mu_{i,p} \cdot \log[(1-s)\mu_i + sx_i]| \\ &= -(1-s) \sum_{i=0}^m |\mu_{i,p}| \log[(1-s)\mu_i + sx_i] \\ &\leq -(1-s) \sum_{i=0}^m |\mu_{i,p}| \log[(1-s)\mu_i], \end{aligned}$$

where the last inequality is due to the inequality  $(1-s)\mu_i + sx_i \geq (1-s)\mu_i$  and the fact that  $\log(\cdot)$  is an increasing function. Next, we note that there exists a neighborhood  $N$  of  $(1, \tilde{\rho}, \tilde{x})$  such that for all  $(s, \rho, x)$  in  $N$  and all  $i = 0, 1, \dots, m$  we have  $|\mu_{i,p}| < M$  and  $-\log[\mu_i] < M$  for some  $M > 0$ . Letting  $\varepsilon > 0$ , there also exists a neighborhood  $\tilde{N}$  of  $(1, \tilde{\rho}, \tilde{x})$  such that  $-(1-s)\log[(1-s)] < \varepsilon$  and  $(1-s) < \varepsilon$  for all  $(s, \rho, x)$  in  $\tilde{N}$ . It follows that the summation  $\sum_{i=0}^m (1-s)\mu_{i,p} \cdot \log[(1-s)\mu_i + sx_i]$  tends to 0 as  $(s, \rho, x)$  tends to  $(1, \tilde{\rho}, \tilde{x})$ . Therefore, we obtain

$$\lim_{(s, \rho, x) \rightarrow (1, \tilde{\rho}, \tilde{x})} B_{\rho_p}(s, \rho, x) = -[\tilde{x}_p - \mu_p(\tilde{\rho})]. \quad (3.4.8)$$

## CHAPTER 3. NOTATION, FUNCTIONS, AND MOTIVATION

Finally, for the limit of  $B_{\rho_p \rho_q}$ , we note that the term  $(2s-1)\mu_{p,q}$  tends to  $\mu_{p,q}(\tilde{\rho})$  as  $(s, \rho, x)$  tends to  $(1, \tilde{\rho}, \tilde{x})$  and the summation  $(1-s) \sum_{i=0}^m [\log[(1-s)\mu_i + sx_i] - \rho_i] \mu_{i,p,q}$  tends to 0 for the same reason that  $\sum_{i=0}^m (1-s) [\log[(1-s)\mu_i + sx_i] - \rho_i] m_{i,p}$  tends to 0 in the proof of (3.4.8). We are left to investigate the limit of the remaining summand  $\sum_{i=0}^m \mu_{i,q} \mu_{i,p} \cdot (1-s)^2 [(1-s)\mu_i + sx_i]^{-1}$  as  $(s, \rho, x)$  tends to  $(1, \tilde{\rho}, \tilde{x})$ . First taking the absolute value and applying the triangle inequality, we have

$$\begin{aligned} \left| \sum_{i=0}^m \frac{\mu_{i,q} \mu_{i,p} \cdot (1-s)^2}{(1-s)\mu_i + sx_i} \right| &\leq \sum_{i=0}^m \frac{|\mu_{i,q} \mu_{i,p}| (1-s)^2}{(1-s)\mu_i + sx_i} \\ &\leq \sum_{i=0}^m \frac{|\mu_{i,q} \mu_{i,p}| (1-s)^2}{(1-s)\mu_i} \\ &= (1-s) \sum_{i=0}^m \frac{|\mu_{i,q} \mu_{i,p}|}{\mu_i}. \end{aligned}$$

Thus, the summand tends to 0 as  $(s, \rho, x)$  tends to  $(1, \tilde{\rho}, \tilde{x})$ , because  $\sum_{i=0}^m |\mu_{i,q} \mu_{i,p}| \mu_i^{-1}$  is uniformly bounded in a small enough neighborhood of  $(1, \tilde{\rho}, \tilde{x})$ . We conclude that

$$\lim_{(s, \rho, x) \rightarrow (1, \tilde{\rho}, \tilde{x})} B_{\rho_p \rho_q}(s, \rho, x) = \mu_{p,q}(\tilde{\rho}).$$

### 3.4.3 Summary for $B$

Here we summarize the computations for  $B$  and its partial derivatives. Throughout this summary, we write  $f(1, \tilde{\rho}, \tilde{x})$  to mean  $\lim_{(s, \rho, x) \rightarrow (1, \tilde{\rho}, \tilde{x})} f(s, \rho, x)$ .

The function

$$B(s, \rho, x) = b(\Lambda(s, \rho, x), \rho)$$

### CHAPTER 3. NOTATION, FUNCTIONS, AND MOTIVATION

is continuous on  $[0, 1] \times \mathbb{R}^m \times \partial\Delta$  and we have

$$B(0, \rho, x) = 0, \text{ and } B(1, \rho, x) = b(x, \rho).$$

Taking a partial derivative of  $B$  with respect to  $s$ , the function  $B_s$  is only continuous on  $[0, 1) \times \mathbb{R}^m \times \partial\Delta$  and we have

$$B_s(0, \rho, x) = 0 \text{ and } B_s(1, \rho, x) = \infty.$$

The function  $B_{ss}$ , like  $B_s$ , is only continuous on  $[0, 1) \times \mathbb{R}^m \times \partial\Delta$  and we have

$$B_{ss}(0, \rho, x) = \sum_{i=0}^m \frac{(x_i - \mu_i)^2}{\mu_i} \text{ and } B_{ss}(1, \rho, x) = \infty.$$

In contrast, the function  $B_{\rho_p}$  can be extended continuously to  $[0, 1] \times \mathbb{R}^m \times \partial\Delta$  (from  $[0, 1) \times \mathbb{R}^m \times \partial\Delta$ ) and we have

$$B_{\rho_p}(0, \rho, x) = 0 \text{ and } B_{\rho_p}(1, \rho, x) = -[x_p - \mu_p(\rho)].$$

Finally, the function  $B_{\rho_p \rho_q}$  can also be extended continuously to  $[0, 1] \times \mathbb{R}^m \times \partial\Delta$  (from  $[0, 1) \times \mathbb{R}^m \times \partial\Delta$ ) and we have

$$B_{\rho_p \rho_q}(0, \rho, x) = 0 \text{ and } B_{\rho_p \rho_q}(1, \rho, x) = \mu_{p,q}(p).$$

## 3.5 Properties of the function $h$

Having detailed the essential computations for the partial derivatives of  $B$ , we return to the function  $h$  and recall that

$$h : H \rightarrow (0, 1)$$

$$(t, \rho, x) \mapsto s \text{ such that } B(s, \rho, x) = t,$$

where  $H = \{(t, \rho, x) \in (0, \infty) \times \mathbb{R}^m \times \partial\Delta \mid 0 < t < b(x, \rho)\}$ . Differentiating the equation  $B(h(t, \rho, x), \rho, x) = t$  implicitly in  $t$  and  $\rho_p$ , we obtain

$$h_t(t, \rho, x) = \frac{1}{B_s(h, \rho, x)} > 0, \quad (3.5.1)$$

and

$$h_{\rho_p}(t, \rho, x) = -\frac{B_{\rho_p}(h, \rho, x)}{B_s(h, \rho, x)}. \quad (3.5.2)$$

We recall these partial derivatives are continuous on  $H$  and look to investigate their behavior on the boundary of  $H$ . In section 3.5.1, we will look at the behavior as  $(t, \rho, x)$  tends to  $(0, \tilde{\rho}, \tilde{x})$ ; in section 3.5.2, we will look at the behavior as  $(t, \rho, x)$  tends to  $(b(\tilde{x}, \tilde{\rho}), \tilde{\rho}, \tilde{x})$ ; and in section 3.5.3, we will summarize these computations.

### 3.5.1 Along $(0, \rho, x)$ .

Fix  $(\tilde{\rho}, \tilde{x}) \in \mathbb{R}^n \times \partial\Delta$ . Also consider the points  $(t, \rho, x) \in H$ .

For the limit of  $h$  as  $(t, \rho, x)$  tends to  $(0, \tilde{\rho}, \tilde{x})$ , we keep in mind that  $h$  is an

### CHAPTER 3. NOTATION, FUNCTIONS, AND MOTIVATION

increasing function with respect to  $t$ . Let  $\varepsilon > 0$ , then there exists  $\delta_1 > 0$  such that  $|h(\delta_1, \tilde{\rho}, \tilde{x})| < \varepsilon/2$  and  $\delta_1 < \tilde{t}/2$ . By the continuity of  $h$  on the interior of  $H$ , there exists  $\delta_2 > 0$  such that  $|h(\delta_1, \tilde{\rho}, \tilde{x}) - h(\delta_1, \rho, x)| < \varepsilon/2$  for all  $\|(\rho, x) - (\tilde{\rho}, \tilde{x})\| < \delta_2$ . Then let  $\delta = \min(\delta_1, \delta_2)$  so that  $\|(t, \rho, x) - (0, \tilde{\rho}, \tilde{x})\| < \delta$  implies

$$\begin{aligned} 0 &\leq h(t, \rho, x) < h(\delta, \rho, x) \\ &\leq h(\delta_1, \rho, x) \\ &= h(\delta_1, \rho, x) - h(\delta_1, \tilde{\rho}, \tilde{x}) + h(\delta_1, \tilde{\rho}, \tilde{x}) \\ &< \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon. \end{aligned}$$

Thus we have

$$\lim_{(t, \rho, x) \rightarrow (0, \tilde{\rho}, \tilde{x})} h(t, \rho, x) = 0. \quad (3.5.3)$$

For the next limit, we look at (3.5.1) in conjunction with (3.4.2) and (3.5.3) and immediately obtain

$$\lim_{(t, \rho, x) \rightarrow (0, \tilde{\rho}, \tilde{x})} h_t(t, \rho, x) = \infty.$$

For the limit of  $h_{\rho_p}$  as  $(t, \rho, x)$  tends to  $(0, \tilde{\rho}, \tilde{x})$ , we apply the mean value theorem and imitate the idea of l'Hospital's rule. By the mean value theorem, we have

$$\begin{aligned} h_{\rho_p}(t, \rho, x) &= -\frac{B_{\rho_p}(h, \rho, x)}{B_s(h, \rho, x)} \cdot \frac{\frac{1}{h}}{\frac{1}{h}} \\ &= -\frac{B_{\rho_p s}(\xi_1, \rho, x)}{B_{ss}(\xi_2, \rho, x)} \end{aligned}$$



## CHAPTER 3. NOTATION, FUNCTIONS, AND MOTIVATION

for some  $\xi_1, \xi_2 \in (0, h)$ . Together with (3.4.4), (3.4.6), and (3.5.3), we obtain

$$\lim_{(t, \rho, x) \rightarrow (0, \tilde{\rho}, \tilde{x})} h_{\rho_p}(t, \rho, x) = 0. \quad (3.5.4)$$

Finally, we look at the limit of  $h \cdot h_t$  as  $(t, \rho, x)$  tends to  $(0, \tilde{\rho}, \tilde{x})$ . As with the proof of 3.5.4, we apply the mean value theorem and imitate the idea of l'Hospital's rule. By the mean value theorem, we have

$$\begin{aligned} h(t, \rho, x)h_t(t, \rho, x) &\stackrel{3.5.1}{=} \frac{1}{\frac{1}{h}B_s(h, \rho, x)} \\ &= \frac{1}{B_{ss}(\xi, \rho, x)} \end{aligned}$$

for some  $\xi \in (0, h)$ . By a combination of (3.5.3) and (3.4.4), we have

$$\lim_{(t, \rho, x) \rightarrow (0, \tilde{\rho}, \tilde{x})} h(t, \rho, x)h_t(t, \rho, x) = \left[ \sum_{i=0}^m \frac{[x_i - \mu_i(\tilde{\rho})]^2}{\mu_i(\tilde{\rho})} \right]^{-1}.$$

Thus, while  $h_t$  tends to  $\infty$  as  $(t, \rho, x)$  tends to  $(0, \tilde{\rho}, \tilde{x})$ , the product of  $h$  with  $h_t$  has a finite limit as  $(t, \rho, x)$  tends to  $(0, \tilde{\rho}, \tilde{x})$ .

### 3.5.2 Along $(b(x, \rho), \rho, x)$ .

Fix  $(\tilde{\rho}, \tilde{x}) \in \mathbb{R}^n \times \partial\Delta$  and  $\tilde{t} = b(\tilde{x}, \tilde{\rho})$ . Also consider the points  $(t, \rho, x) \in H$ .

Let us start by determining the limit of  $h$  as  $(t, \rho, x)$  tends to  $(\tilde{t}, \tilde{\rho}, \tilde{x})$ . As with the proof of (3.5.3), we keep in mind that  $h$  is an increasing function with respect to  $t$ .

Let  $\varepsilon > 0$ , then there exists  $\delta_1 > 0$  such that  $\tilde{t} - \delta_1 > 0$  and  $[1 - h(\tilde{t} - \delta_1, \tilde{\rho}, \tilde{x})] < \varepsilon/2$ .

We also have  $\tilde{t} - \delta_1 < \tilde{t} = b(\tilde{x}, \tilde{\rho})$  and choose, by the continuity of  $b$ , a  $\delta_2 > 0$

### CHAPTER 3. NOTATION, FUNCTIONS, AND MOTIVATION

so that  $|b(x, \rho) - b(\tilde{x}, \tilde{\rho})| < \delta_1$  for all  $\|(\rho, x) - (\tilde{\rho}, \tilde{x})\| < \delta_2$ . This ensures  $b(x, \rho) > b(\tilde{x}, \tilde{\rho}) - \delta_1 = \tilde{t} - \delta_1$  for all  $\|(t, \rho, x) - (\tilde{t}, \tilde{\rho}, \tilde{x})\| < \delta_2$  and consequently  $(\tilde{t} - \delta_1, \rho, x) \in H$ . Then by the continuity of  $h$  on  $H$ , there exists  $\delta_3 > 0$  (chosen to be less than  $\delta_2$ ) such that  $|h(\tilde{t} - \delta_1, \tilde{\rho}, \tilde{x}) - h(\tilde{t} - \delta_1, \rho, x)| < \varepsilon/2$  for all  $\|(\rho, x) - (\tilde{\rho}, \tilde{x})\| < \delta_3$ . Then let  $\delta = \min(\delta_1, \delta_3)$  so that  $\|(t, \rho, x) - (\tilde{t}, \tilde{\rho}, \tilde{x})\| < \delta$  implies

$$\begin{aligned} 0 &\leq 1 - h(t, \rho, x) < 1 - h(\tilde{t} - \delta, \rho, x) \\ &\leq 1 - h(\tilde{t} - \delta_1, \rho, x) \\ &= 1 - h(\tilde{t} - \delta_1, \tilde{\rho}, \tilde{x}) + h(\tilde{t} - \delta_1, \tilde{\rho}, \tilde{x}) - h(\tilde{t} - \delta_1, \rho, x) \\ &< \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon. \end{aligned}$$

Thus we have

$$\lim_{(t, \rho, x) \rightarrow (\tilde{t}, \tilde{\rho}, \tilde{x})} h(t, \rho, x) = 1. \quad (3.5.5)$$

Next, we note that (3.4.7) and (3.5.5) imply

$$\lim_{(t, \rho, x) \rightarrow (\tilde{t}, \tilde{\rho}, \tilde{x})} h_t(t, \rho, x) = 0.$$

Finally, we combine (3.4.7), (3.4.8), and (3.5.5) with (3.5.2) to conclude

$$\lim_{(t, \rho, x) \rightarrow (\tilde{t}, \tilde{\rho}, \tilde{x})} h_{\rho_p}(t, \rho, x) = 0.$$

### 3.5.3 Summary for $h$

From the work shown above, we conclude that  $h$  can be extended continuously to  $[0, \infty) \times \mathbb{R}^m \times \partial\Delta$  (from  $(0, b(x, \rho)) \times \mathbb{R}^m \times \partial\Delta$ ). Explicitly, we have

$$h(t, \rho, x) = \begin{cases} 0, & \text{if } t = 0 \\ B(h, \rho, x) = t, & \text{if } 0 \leq t \leq b(x, \rho) \\ 1, & \text{if } t \geq b(x, \rho). \end{cases}$$

The partial derivative  $h_t$  can only be extended continuously to  $(0, \infty) \times \mathbb{R}^n \times \partial\Delta$  (from  $(0, b(x, \rho)) \times \mathbb{R}^n \times \partial\Delta$ ) and it tends towards infinity as  $(t, \rho, x)$  tends to  $(0, \tilde{\rho}, \tilde{x})$  for any  $(\tilde{\rho}, \tilde{x}) \in \mathbb{R}^m \times \partial\Delta$ . We have

$$h_t(t, \rho, x) = \begin{cases} \frac{1}{B_s(h, \rho, x)}, & \text{if } 0 < t \leq b(x, \rho) \\ 0, & \text{if } t \geq b(x, \rho). \end{cases}$$

Of slightly greater importance, we can continuously extend the product  $h \cdot h_t$  to  $[0, \infty) \times \mathbb{R}^m \times \partial\Delta$  (from  $(0, b(x, \rho)) \times \mathbb{R}^m \times \partial\Delta$ ) and obtain

$$(h \cdot h_t)(t, \rho, x) = \begin{cases} \left[ \sum_{i=0}^n \frac{[x_i - \mu_i]^2}{\mu_i} \right]^{-1}, & \text{if } t = 0 \\ \frac{h}{B_s(h, \rho, x)}, & \text{if } 0 \leq t \leq b(x, \rho) \\ 0, & \text{if } t \geq b(x, \rho). \end{cases}$$

Finally, the partial derivative  $h_{\rho_p}$  can be extended continuously to  $[0, \infty) \times \mathbb{R}^m \times \partial\Delta$

### CHAPTER 3. NOTATION, FUNCTIONS, AND MOTIVATION

(from  $(0, b(x, \rho)) \times \mathbb{R}^m \times \partial\Delta$ ). Explicitly, we have

$$h_{\rho_p}(t, \rho, x) = \begin{cases} 0, & \text{if } t = 0 \\ -\frac{B_{\rho_p}(h, \rho, x)}{B_s(h, \rho, x)}, & \text{if } 0 \leq t \leq b(x, \rho) \\ 0, & \text{if } t \geq b(x, \rho). \end{cases}$$

# Chapter 4

## Proof of the Theorem

By (2.2.5), the proof of Theorem 4 reduces to showing

$$G_n(\rho) \doteq \int_0^\infty [1 - D(t, \rho)]^n dt$$

is twice continuously differentiable on  $\mathbb{R}^m$ . Working with only fixed  $n$  and wishing to use subscript notation when taking partial derivatives, we henceforth omit the subscript  $n$  and simply write  $G$  instead of  $G_n$ .

In order to differentiate  $G$ , we first look to express  $D$  as an integral (Lemma 7 in Section 4.1). The proof of Theorem 4 proceeds and concludes by qualifying various applications of Lebesgue's dominated convergence theorem.

## 4.1 Rewriting $D(t, \rho)$ as an integral

Before expressing  $D$  as an integral, let us consider the volume form  $d\lambda = m!d\lambda_1 \wedge \cdots \wedge d\lambda_m$  on the unit simplex  $\Delta$ . Fix  $\rho \in \mathbb{R}^m$  and for  $i = 0, 1, \dots, m$ , we define  $R_i(\rho) \subset \Delta$  as the region between the point  $\mu(\rho) \in \Delta^\circ$  and the  $i$ -th facet  $\partial^i \Delta$  of  $\Delta$ . Explicitly, the set

$$R_i(\rho) \doteq \{\lambda \in \Delta \mid \lambda = (1-s)\mu(\rho) + sx, \text{ for some } x \in \partial^i \Delta \text{ and } 0 \leq s \leq 1\}. \quad (4.1.1)$$

Together with the above notation, we consider  $d\lambda$  under a change of variables given by the map  $\Lambda$  from Section 3.

**Lemma 6.** *Define the map*

$$\Lambda : [0, 1] \times \mathbb{R}^m \times \partial\Delta \twoheadrightarrow \Delta$$

$$(s, \rho, x) \mapsto (1-s)\mu(\rho) + sx$$

and fix a  $\rho \in \mathbb{R}^m$ . Then the function  $\lambda(s, x) = \Lambda(s, \rho, x)$  restricted to  $[0, 1] \times \partial^i \Delta$  is one-to-one and onto  $R_i(\rho) \subset \Delta$ . We also have

$$d\lambda|_{R_i} = \mu_i \cdot m s^{m-1} ds \wedge dx_{(i)} \quad (4.1.2)$$

where  $dx_{(i)} \doteq -(m-1)! \left( \frac{\partial}{\partial x_i} \lrcorner dx \right)$  and  $\frac{\partial}{\partial x_0} \doteq -\frac{1}{m} \sum_{i=1}^m \frac{\partial}{\partial x_i}$ ; geometrically, the vector  $\frac{\partial}{\partial x_i}$  is the inward facing normal to  $\partial^i \Delta$  for  $i = 0, 1, \dots, m$ .

*Proof.* The first part of the lemma is immediate from the surjectivity of  $\Lambda$  onto  $\Delta$ , the convexity of  $\Delta$ , and the definition of  $R_i(\rho)$ . For the second part, let  $\lambda(s, x) =$

## CHAPTER 4. PROOF OF THE THEOREM

$(1-s)\mu + sx$  so that  $d\lambda_i = (x_i - \mu_i)ds + sdx_i$  and

$$\begin{aligned} d\lambda_1 \wedge \cdots \wedge d\lambda_m &= \sum_{i=1}^m (x_i - \mu_i) s^{m-1} dx_1 \wedge \cdots \wedge dx_{i-1} \wedge ds \wedge dx_{i+1} \wedge \cdots \wedge dx_m \\ &= \sum_{i=1}^m (x_i - \mu_i) s^{m-1} (-1)^{i-1} ds \wedge dx_1 \wedge \cdots \widehat{dx_i} \cdots \wedge dx_m \\ &= \sum_{i=1}^m (x_i - \mu_i) s^{m-1} ds \wedge \left( \frac{\partial}{\partial x_i} \lrcorner dx \right). \end{aligned}$$

Fix  $k \in \{1, \dots, m\}$ . Then for any  $x \in \partial^k \Delta$  we have  $x_k = 0$  and this implies  $dx_k = 0$ .

In this case,

$$\sum_{i=1}^m (x_i - \mu_i) s^{m-1} ds \wedge \left( \frac{\partial}{\partial x_i} \lrcorner dx \right) = -\mu_k \cdot s^{m-1} ds \wedge \left( \frac{\partial}{\partial x_k} \lrcorner dx \right)$$

as claimed above. Now let  $k = 0$ . Then for any  $x \in \partial^0 \Delta$  we have  $\sum_{i=1}^m x_i = 1$  and

this implies  $\sum_{i=1}^m dx_i = 0$ . From this we obtain

$$\begin{aligned} \frac{\partial}{\partial x_j} \lrcorner dx &= (-1)^{j-1} dx_1 \wedge \cdots \widehat{dx_j} \cdots \wedge dx_{m-1} \wedge dx_m \\ &= (-1)^{j-1} dx_1 \wedge \cdots \widehat{dx_j} \cdots \wedge dx_{m-1} \wedge \left( -\sum_{i=1}^{m-1} dx_i \right) \\ &= (-1)^j dx_1 \wedge \cdots \widehat{dx_j} \cdots \wedge dx_{m-1} \wedge (dx_j) \\ &= (-1)^j (-1)^{m-j-1} dx_1 \wedge \cdots \wedge dx_{m-1} \\ &= \frac{\partial}{\partial x_m} \lrcorner dx \end{aligned}$$

for any  $j = 1, \dots, m-1$  and hence

$$-\frac{\partial}{\partial x_0} \lrcorner dx = \frac{\partial}{\partial x_j} \lrcorner dx$$

## CHAPTER 4. PROOF OF THE THEOREM

for any  $j = 1, \dots, m$ . Thus we have

$$\begin{aligned} \sum_{i=1}^m (x_i - \mu_i) s^{m-1} ds \wedge \left( \frac{\partial}{\partial x_i} \lrcorner dx \right) &= \sum_{i=1}^m (x_i - \mu_i) s^{m-1} ds \wedge \left( -\frac{\partial}{\partial x_0} \lrcorner dx \right) \\ &= -(x_0 - \mu_0) s^{m-1} ds \wedge \left( -\frac{\partial}{\partial x_0} \lrcorner dx \right) \\ &= -\mu_0 \cdot s^{m-1} ds \wedge \left( \frac{\partial}{\partial x_0} \lrcorner dx \right) \end{aligned}$$

as claimed above. For an alternative proof, rewrite  $d\lambda_m = (x_m - \mu_m)ds + sdx_m =$

$(x_m - \mu_m)ds - sdx_1 - \dots - sdx_{m-1}$  and so  $d\lambda_1 \wedge \dots \wedge d\lambda_m = \det(M)ds \wedge dx_1 \wedge \dots \wedge dx_{m-1}$

where

$$M = \begin{bmatrix} (x_1 - \mu_1) & s & 0 & \dots & 0 \\ (x_2 - \mu_2) & 0 & s & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ (x_{m-1} - \mu_{m-1}) & 0 & 0 & \dots & s \\ (x_m - \mu_m) & -s & -s & \dots & -s \end{bmatrix}.$$



## CHAPTER 4. PROOF OF THE THEOREM

By switching rows and row reducing, we obtain

$$\begin{aligned}
 \det M &= (-1)^{m-1} \det \begin{bmatrix} (x_m - \mu_m) & -s & -s & \cdots & -s \\ (x_1 - \mu_1) & s & 0 & \cdots & 0 \\ (x_2 - \mu_2) & 0 & s & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ (x_{m-1} - \mu_{m-1}) & 0 & 0 & \cdots & s \end{bmatrix} \\
 &= (-1)^{m-1} \det \begin{bmatrix} (\sum_{i=1}^m x_i - \sum_{i=1}^m \mu_i) & 0 & 0 & \cdots & 0 \\ (x_1 - \mu_1) & s & 0 & \cdots & 0 \\ (x_2 - \mu_2) & 0 & s & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ (x_{m-1} - \mu_{m-1}) & 0 & 0 & \cdots & s \end{bmatrix} \\
 &= (-1)^{m-1} \mu_0 s^{n-1}.
 \end{aligned}$$

□

Having proven Lemma 6, we are ready to take the vague notion of writing  $D$  as an integral given in Section 3.2 and make it rigorous.

**Lemma 7.** *The function  $D(t, \rho) = \mathcal{P}\{\lambda \in \Delta \mid b(\lambda, \rho) \leq t\}$  can be expressed as an integral. We have*

$$D(t, \rho) = \sum_{i=0}^m \mu_i \cdot \int_{\partial^i \Delta} h^m(t, \rho, x) dx_{(i)} \quad (4.1.3)$$

where  $dx_{(i)} \doteq -(n-1)! \left( \frac{\partial}{\partial x_i} \lrcorner dx \right)$  and  $\frac{\partial}{\partial x_0} \doteq -\frac{1}{m} \sum_{i=1}^m \frac{\partial}{\partial x_i}$ . Moreover, the function  $D$  is continuous on  $[0, \infty) \times \mathbb{R}^m$ .

## CHAPTER 4. PROOF OF THE THEOREM

*Proof.* Let  $S = S(t, \rho) \doteq \{\lambda \in \Delta \mid b(\lambda, \rho) \leq t\}$  so that  $D = \int_S d\lambda$  and define  $S_i \doteq S \cap R_i$ , where  $R_i$  are defined at (4.1.1). It follows immediately that  $S = \cup_{i=0}^m S_i$  and  $\mathcal{P}(S_i \cap S_j) = 0$  for  $i \neq j$ . Thus far, we have

$$D(t, \rho) = \sum_{i=0}^m \int_{S_i(t, \rho)} d\lambda$$

and proceed by giving  $S_i(t, \rho)$  explicitly. We claim that

$$S_i(t, \rho) = \{\lambda \in \Delta \mid \lambda = (1-s)\mu(\rho) + sx \text{ for some } x \in \partial^i \Delta \text{ and } 0 \leq s \leq h(t, \rho, x)\}. \quad (4.1.4)$$

Starting with the forward inclusion, we take  $\lambda \in S_i = S \cap R_i$ . Then  $b(\lambda, \rho) \leq t$  and  $\lambda = (1-s)\mu(\rho) + sx$  for some  $x \in \partial^i \Delta$  and some  $s \in [0, 1]$ . These conditions imply that  $B(s, \rho, x) = b(\lambda, \rho) \leq t$ . Then since  $h(\cdot, \rho, x)$  is an increasing function and inverse to  $B(\cdot, \rho, x)$ , we obtain  $s = h(B(s, \rho, x), \rho, x) \leq h(t, \rho, x)$  and this gives the desired inclusion. For the reverse inclusion, let  $\lambda = (1-s)\mu(\rho) + sx$  for some  $x \in \partial^i \Delta$  and  $0 \leq s \leq h(t, \rho, x)$ . Then because  $B(\cdot, \rho, x)$  is an increasing function and inverse to  $h(\cdot, \rho, x)$  we have  $b(\lambda, \rho) = B(s, \rho, x) \leq B(h(t, \rho, x), \rho, x) = t$ . This concludes the reverse inclusion and proves (4.1.4).

We next make the change of variables  $\lambda(s, x) = (1-s)\mu(\rho) + sx$ . Using Lemma

## CHAPTER 4. PROOF OF THE THEOREM

6, we obtain

$$\begin{aligned}
 \sum_{i=0}^m \int_{S_i(t,\rho)} d\lambda &\stackrel{(4.1.2)}{=} \sum_{i=0}^m \int_{S_i(t,\rho)} \mu_i \cdot m s^{m-1} ds \wedge dx_{(i)} \\
 &= \sum_{i=0}^m \mu_i \cdot \int_{\partial^i \Delta} \left[ \int_0^{h(t,\rho,x)} m s^{m-1} ds \right] dx_{(i)} \\
 &= \sum_{i=0}^m \mu_i \cdot \int_{\partial^i \Delta} h^m(t, \rho, x) dx_{(i)}.
 \end{aligned}$$

Note that we can apply Fubini's theorem in the second equality, because  $0 \leq D \leq 1$  and hence the integral is finite.

The continuity of  $D$  follows from Lebesgue's dominated convergence theorem.  $\square$

*Remark 8.* For  $m = 1$ , the formulation (4.1.3) is the integration of a 0-form over a 0-chain, which is evaluation at a point. Explicitly, we have

$$\begin{aligned}
 D(t, \rho) &= \sum_{i=0}^1 \mu_i \cdot \int_{\partial^i \Delta} h(t, \rho, x) dx_{(i)} \\
 &= \sum_{i=0}^1 \mu_i \cdot h(t, \rho, \partial^i \Delta) \\
 &= \mu_0 \cdot h(t, \rho, \partial^0 \Delta) + \mu_1 \cdot h(t, \rho, \partial^1 \Delta) \\
 &= \mu_0 \cdot h(t, \rho, 1) + \mu_1 \cdot h(t, \rho, 0),
 \end{aligned} \tag{4.1.5}$$

which agrees with (3.2.5).

## 4.2 The first partial derivative of $G$

Now we look to differentiate our function  $G$ , bring the derivative under the integral sign, and conclude the derivative is continuous.

**Lemma 9.** *Given the function*

$$D(t, \rho) = \sum_{i=0}^m \mu_i \cdot \int_{\partial^i \Delta} h^m(t, \rho, x) dx_{(i)}$$

*on  $[0, \infty) \times \mathbb{R}^m$ , the partial derivative  $D_{\rho_p}$  exists and is continuous on  $[0, \infty) \times \mathbb{R}^m$ .*

*It is given by*

$$D_{\rho_p} = \sum_{i=0}^m \mu_i \cdot \int_{\partial^i \Delta} m h^{m-1} h_{\rho_p} dx_{(i)} + \mu_{i,p} \cdot \int_{\partial^i \Delta} h^m dx_{(i)}.$$

*Consequently, the partial derivative of*

$$G(\rho) = \int_0^\infty [1 - D(t, \rho)]^n dt$$

*exists and is continuous on  $\mathbb{R}^m$ . It is given by*

$$G_{\rho_p} = - \int_0^\infty n[1 - D]^{n-1} D_{\rho_p} dt.$$

*Proof.* Referring to Section 3.5.3, the function  $h^m$  and its partial derivative with respect to  $\rho_p$  are continuous on  $[0, \infty) \times \mathbb{R}^m \times \partial\Delta$ . Thus, by Lebesgue's dominated convergence theorem, we may differentiate under the integral sign and obtain

$$\frac{\partial}{\partial \rho_p} \int_{\partial^i \Delta} h^m dx_{(i)} = \int_{\partial^i \Delta} m h^{m-1} h_{\rho_p} dx_{(i)}.$$

## CHAPTER 4. PROOF OF THE THEOREM

We note that the dominated convergence theorem also implies the right-hand side is continuous on  $[0, \infty) \times \mathbb{R}^m$ .

The function  $D$  is continuous on  $[0, \infty) \times \mathbb{R}^m$  (Lemma 7) and the function  $D_{\rho_p}$  is continuous on  $[0, \infty) \times \mathbb{R}^m$  (first part of this lemma). Thus, the function  $\frac{\partial}{\partial \rho_p}[1 - D]^n$  exists, is given by  $n[1 - D]^{n-1}D_{\rho_p}$ , and is continuous on  $[0, \infty) \times \mathbb{R}^m$ . Next we let  $\tilde{\rho} \in \mathbb{R}^m$ , choose  $r > 0$ , and consider the support of the family of functions  $t \mapsto 1 - D(t, \rho)$  with  $\rho \in B_r(\tilde{\rho})$ . For a fixed  $\rho \in B_r(\tilde{\rho})$ , the support of  $t \mapsto 1 - D(t, \rho)$  is contained in the interval  $[0, \max_{\lambda \in \Delta} b(\lambda, \rho)]$ . We recalling from Lemma 5 that  $\max_{\lambda \in \Delta} b(\lambda, \rho) = \max_{i \in \{0, 1, \dots, m\}} \{\log(1 + |e^\rho|) - \rho_i\}$ . For  $i = 0, 1, \dots, m$ , there exists  $M_i > 0$ , depending on  $\tilde{\rho}$  and  $r$ , such that  $\log(1 + |e^\rho|) - \rho_i < M_i$  for all  $\rho \in B_r(\tilde{\rho})$ . Consequently, there exists an  $M > 0$ , depending on  $\tilde{\rho}$  and  $r$ , such that  $\text{supp}(1 - D(\cdot, \rho)) \subset [0, M]$  for all  $\rho \in B_r(\tilde{\rho})$ . By Lebesgue's dominated convergence theorem, we may differentiate under the integral sign and obtain

$$G_{\rho_p} = \frac{\partial}{\partial \rho_p} \int_0^\infty [1 - D]^n dt = - \int_0^\infty n[1 - D]^{n-1} D_{\rho_p} dt.$$

The theorem also implies the function is continuous on  $\mathbb{R}^m$ . □

### 4.3 The second partial derivative of $G$

In Lemma 9 we found that

$$D_{\rho_p} = \sum_{i=0}^m \mu_i \cdot \int_{\partial^i \Delta} m h^{m-1} h_{\rho_p} dx_{(i)} + \mu_{i,p} \cdot \int_{\partial^i \Delta} h^m dx_{(i)}$$

## CHAPTER 4. PROOF OF THE THEOREM

is continuous on  $[0, \infty) \times \mathbb{R}^m$ . Then let  $S = \sum_{i=0}^m \mu_i \cdot \int_{\partial^i \Delta} m h^{m-1} h_{\rho_p} dx_{(i)}$  and  $\tilde{S} = \mu_{i,p} \cdot \int_{\partial^i \Delta} h^m dx_{(i)}$ . Thus, we have

$$G_{\rho_p} = - \int_0^\infty n[1 - D]^{n-1} S dt - \int_0^\infty n[1 - D]^{n-1} \tilde{S} dt.$$

We first proceed with the second summand  $-\int_0^\infty n[1 - D]^{n-1} \tilde{S} dt$ , because it is easier to deal with, and then look at the first summand  $-\int_0^\infty n[1 - D]^{n-1} S dt$

### 4.3.1 The term $-\int_0^\infty n[1 - D]^{n-1} \tilde{S} dt$ is continuously differentiable

The factor  $\tilde{S}$  is continuous and has continuous derivatives in  $\rho_p$  over  $[0, \infty) \times \mathbb{R}^m$  for the same reason that  $D$  has continuous derivatives in  $\rho_p$  over  $[0, \infty) \times \mathbb{R}^m$  (Lemma 7, Lemma 9). Then the integrand of the second summand  $n[1 - D]^{n-1} \tilde{S}$  and its derivative in  $\rho_q$ ,  $n(n-1)[1 - D]^{n-2} D_{\rho_q} \tilde{S} + n[1 - D]^{n-1} \tilde{S}_{\rho_q}$ , are continuous over  $[0, \infty) \times \mathbb{R}^m$ . Imitating the proof of Lemma 9, the first partial derivative of the second summand with respect to  $\rho_q$  exists, is given by

$$\frac{\partial}{\partial \rho_q} \int_0^\infty n[1 - D]^{n-1} \tilde{S} dt = \int_0^\infty \frac{\partial}{\partial \rho_q} [n[1 - D]^{n-1} \tilde{S}] dt,$$

and is continuous over  $\mathbb{R}^m$ .

### 4.3.2 The term $-\int_0^\infty n[1-D]^{n-1}Sdt$ is continuously differentiable

Now we are left to see that  $\int_0^\infty n[1-D]^{n-1}Sdt$  has a continuous partial derivative.

We first write

$$\begin{aligned} \int_0^\infty n[1-D]^{n-1}Sdt &= \int_0^\infty n[1-D]^{n-1} \left[ \sum_{i=0}^m \mu_i \cdot \int_{\partial^i \Delta} m h^{m-1} h_{\rho_p} dx_{(i)} \right] dt \quad (4.3.1) \\ &= nm \sum_{i=0}^m \mu_i \cdot \int_0^\infty \left[ \int_{\partial^i \Delta} [1-D]^{n-1} h^{m-1} h_{\rho_p} dx_{(i)} \right] dt \\ &= nm \sum_{i=0}^m \mu_i \cdot T_i, \end{aligned}$$

where

$$T_i(\rho) \doteq \int_0^\infty \left[ \int_{\partial^i \Delta} [1-D(t, \rho)]^{n-1} h^{m-1}(t, \rho, x) h_{\rho_p}(t, \rho, x) dx_{(i)} \right] dt. \quad (4.3.2)$$

Applying Fubini's theorem and making the change of variables  $t(s) = B(s, \rho, x)$ , we obtain

$$\begin{aligned} T_i &= \int_{\partial^i \Delta} \left[ \int_0^{b(x, \rho)} [1-D]^{n-1} h^{m-1} h_{\rho_p} dt \right] dx_{(i)} \\ &= \int_{\partial^i \Delta} \left[ \int_0^1 [1-D(B, \rho)]^{n-1} s^{m-1} h_{\rho_p}(B, \rho, x) B_s ds \right] dx_{(i)}. \end{aligned}$$

We note that the change of variables is motivated by our desire to replace a moving domain of integration with a fixed domain of integration [Fla73]. Finally, since

## CHAPTER 4. PROOF OF THE THEOREM

$h_{\rho_p}(B, \rho, x)B_s = -B_{\rho_p}$  (3.5.2), we have

$$T_i(\rho) = \int_{\partial^i \Delta} \left[ \int_0^1 s^{m-1} [1 - D(B, \rho)]^{n-1} (-B_{\rho_p}) ds \right] dx_{(i)}. \quad (4.3.3)$$

Section (3.4.3) and Lemma 7 imply the integrand  $s^{m-1}[1 - D(B, \rho)]^{n-1}(-B_{\rho_p})$  is a continuous function on  $[0, 1] \times \mathbb{R}^n \times \partial\Delta$ . It's partial derivative in  $\rho_q$  is

$$s^{m-1}(n-1)[1 - D(B, \rho)]^{n-2}[D_t(B, \rho)B_{\rho_q} + D_{\rho_q}(B, \rho)]B_{\rho_p} - s^{m-1}[1 - D(B, \rho)]^{n-1}B_{\rho_p\rho_q}.$$

Except for  $(s, \rho, x) \mapsto D_t(B(s, \rho, x), \rho)$ , all the functions appearing in (4.3.2) have been shown to be continuous on  $[0, 1] \times \mathbb{R}^m \times \partial\Delta$ . More specifically, the continuity of  $D$  is given by Lemma 7, the continuity of  $D_{\rho_q}$  is given by Lemma 9, and the continuity of  $B$ ,  $B_{\rho_p}$ , and  $B_{\rho_p\rho_k}$  are stated in Section 3.4.3. Thus, in order to show 4.3.2 is continuous on  $[0, 1] \times \mathbb{R}^m \times \partial\Delta$ , we would like to show  $D_t$  is continuous on  $[0, \infty) \times \mathbb{R}^m$ . However, we only have the following lemma:

**Lemma 10.** *For  $m > 1$ , the partial derivative  $D_t$  exists and is continuous on  $[0, \infty) \times \mathbb{R}^m$ . It is given by*

$$D_t = \sum_{i=0}^m \mu_i \cdot \int_{\partial^i \Delta} m h^{m-1} h_t dx_{(i)} \geq 0.$$

*For  $m = 1$ , the partial derivative  $D_t$  exists and is continuous on  $(0, \infty) \times \mathbb{R}$ . It is given by*

$$D_t(t, \rho) = \sum_{i=0}^1 \mu_i \cdot h_t(t, \rho, \partial^i \Delta). \quad (4.3.4)$$

*Proof.* As stated in Section 3.5.3, the functions  $h$  and  $h \cdot h_t$  are continuous on  $[0, \infty) \times$



## CHAPTER 4. PROOF OF THE THEOREM

$\mathbb{R}^m \times \partial\Delta$ . Thus the function  $h^m$  and its partial derivative  $mh^{m-1}h_t$  (for  $m \geq 2$ ) are continuous on  $[0, \infty) \times \mathbb{R}^m \times \partial\Delta$ . By Lebesgue's dominated convergence theorem, we differentiate under the integral sign and obtain

$$\frac{\partial}{\partial t} \int_{\partial^i \Delta} h^m dx_{(i)} = \int_{\partial^i \Delta} mh^{m-1}h_t dx_{(i)}.$$

The theorem also gives the continuity of the function on  $[0, \infty) \times \mathbb{R}^m$ .

While we can take  $m = 1$  as a special case of  $m > 1$  except on  $(0, \infty) \times \mathbb{R}^m \times \partial\Delta$ , we instead choose to differentiate the expression for  $D$  given by (4.1.5) with respect to  $t$  and refer to the continuity of  $h_t$  as stated in Section 3.5.3.  $\square$

Thus, the function  $D_t$  is only continuous on  $[0, \infty) \times \mathbb{R}^m$  for  $m > 1$ . While  $D_t(B(s, \rho, x), \rho)$  tends to  $\infty$  as  $(s, \rho, x)$  tends to  $(0, \tilde{\rho}, \tilde{x})$  for any  $(\tilde{\rho}, \tilde{x}) \in \mathbb{R}^m \times \partial\Delta$ , the function  $B_\rho(s, \rho, x)$  tends to 0 as  $(s, \rho, x)$  tends to  $(0, \tilde{\rho}, \tilde{x})$ . It would be convenient then, for  $m = 1$ , if the map  $(s, \rho, x) \mapsto D_t(B(s, \rho, x), \rho)B_\rho(s, \rho, x)$  could be extended continuously to  $[0, 1] \times \mathbb{R} \times \partial\Delta$ . Since Section (3.4.3) and Lemma 10 imply it is continuous on  $(0, 1] \times \mathbb{R} \times \partial\Delta$ , it suffices to show the limit as  $(s, \rho, x)$  tends to  $(0, \tilde{\rho}, \tilde{x})$  for any  $(\tilde{\rho}, \tilde{x}) \in \mathbb{R} \times \partial\Delta$  exists.

However, before setting out to do this, we shall prove a technical lemma concerning the function  $B$ . The first part of the lemma essentially says that since  $s \mapsto B(s, \rho, 0)$  and  $s \mapsto B(s, \rho, 1)$  are strictly increasing functions starting at 0 for  $s = 0$ , for sufficiently small  $s$ , we can find  $\tilde{s}$  such that  $B(s, \rho, 0) = B(\tilde{s}, \rho, 1)$  and vice versa. The second part of the lemma essentially says that since  $B(\cdot, \rho, 0)$  and  $B(\cdot, \rho, 1)$  are shaped

## CHAPTER 4. PROOF OF THE THEOREM

similarly, the rate at which  $\tilde{s}$  goes to 0 is almost proportional to the rate at which  $s$  goes to 0. Rigorously, we have the following:

**Lemma 11.** *There exists  $\delta > 0$  such that for any  $s \in [0, \delta]$ , there exists  $\tilde{s} = \tilde{s}(s, \rho)$  such that*

$$B(s, \rho, x) = B(\tilde{s}, \rho, 1 - x) \quad (4.3.5)$$

Furthermore, we have

$$\frac{d\tilde{s}}{ds}(0, \rho) = \lim_{s \rightarrow 0} \frac{\tilde{s}}{s} = \exp[(-1)^x \rho].$$

*Proof.* For the first part, we recall from Section 3.2 that the function  $b(\cdot, \rho)$  is convex with a minimum at  $\mu(\rho)$  so its inverse with two branches,  $g_{-1}$  and  $g_1$ , given by

$$g_{-1}(t, \rho) = \lambda \quad \text{if } t \in [0, b(0, \rho)], \quad b(\lambda, \rho) = t, \quad \text{and } \lambda \in [0, \mu(\rho)]$$

and

$$g_1(t, \rho) = \lambda \quad \text{if } t \in [0, b(1, \rho)], \quad b(\lambda, \rho) = t, \quad \text{and } \lambda \in [\mu(\rho), 1].$$

Thus, for  $t \in [0, \min\{b(0, \rho), b(1, \rho)\}]$ , we have  $g_{-1}(t, \rho) \in [0, \mu(\rho)]$  and  $g_1(t, \rho) \in [\mu(\rho), 1]$  with  $b(g_{-1}, \rho) = t = b(g_1, \rho)$ . Separately, given  $\lambda_{-1} \in [0, \mu(\rho)]$  and  $\lambda_1 \in [\mu(\rho), 1]$ , there exists  $s_{-1} \geq 0$  and  $s_1 \geq 0$  such that  $\lambda_{-1} = (1 - s_{-1})\mu + s_{-1} \cdot 0$  and  $\lambda_1 = (1 - s_1)\mu + s_1 \cdot 1$ . Letting  $\varepsilon = \min\{b(0, \rho), b(1, \rho)\}$ , there is a  $\delta > 0$  such that  $0 \leq B(s, \rho, x) < \varepsilon$  for all  $s \in [0, \delta]$ . Then for any  $s \in [0, \delta]$ , we have  $B(s, \rho, 0) \in [0, \min\{b(0, \rho), b(1, \rho)\}]$  and so  $g_1(B(s, \rho, 0), \rho) \in [\mu(\rho), 1]$  and we find

## CHAPTER 4. PROOF OF THE THEOREM

$\tilde{s} \geq 0$  such that  $g_1(B(s, \rho, 0), \rho) = (1 - \tilde{s})\mu + \tilde{s} \cdot 1$ . Taking  $b(\cdot, \rho)$  of both sides, we obtain

$$\begin{aligned} B(s, \rho, 0) &= b((1 - \tilde{s})\mu + \tilde{s} \cdot 1, \rho) \\ &= B(\tilde{s}, \rho, 1). \end{aligned}$$

Likewise, for any  $r \in [0, \delta]$ , we have  $B(r, \rho, 1) \in [0, \min\{b(0, \rho), b(1, \rho)\}]$  and so  $g_{-1}(B(r, \rho, 1), \rho) \in [0, \mu(\rho)]$  and we find  $\tilde{r} \geq 0$  such that  $g_{-1}(B(r, \rho, 1), \rho) = (1 - \tilde{r})\mu + \tilde{r} \cdot 0$ . Taking  $b(\cdot, \rho)$  of both sides, we obtain

$$\begin{aligned} B(r, \rho, 1) &= b((1 - \tilde{r})\mu + \tilde{r} \cdot 0, \rho) \\ &= B(\tilde{r}, \rho, 0) \end{aligned}$$

For the second part, we have  $B(0, \rho, x) = B_s(0, \rho, x) = 0$  and

$$\begin{aligned} B_{ss}(0, \rho, x) &= \frac{(x_0 - \mu_0)^2}{\mu_0} + \frac{(x_1 - \mu_1)^2}{\mu_1} \\ &= \frac{((1 - x) - (1 - \mu))^2}{1 - \mu} + \frac{(x - \mu)^2}{\mu} \\ &= (x - \mu)^2 \frac{1}{\mu(1 - \mu)} \\ &= \begin{cases} e^\rho & \text{if } x = 0 \\ e^{-\rho} & \text{if } x = 1 \end{cases} \end{aligned}$$

Then by Taylor's theorem centered around 0, we have

$$B(s, \rho, 0) = \frac{e^\rho}{2}s^2 + \frac{1}{6}B_{sss}(\xi, \rho, 0)s^3 \quad (4.3.6)$$

## CHAPTER 4. PROOF OF THE THEOREM

for some  $\xi \in (0, s)$  and

$$B(\tilde{s}, \rho, 1) = \frac{e^{-\rho}}{2} \tilde{s}^2 + \frac{1}{6} B_{sss}(\tilde{\xi}, \rho, 1) \tilde{s}^3 \quad (4.3.7)$$

for some  $\tilde{\xi} \in (0, \tilde{s})$ . Combining (4.3.5), (4.3.6), and (4.3.7), we have

$$\left[ \frac{e^\rho}{2} + \frac{1}{6} B_{sss}(\xi, \rho, 0) s \right] s^2 = \left[ \frac{e^{-\rho}}{2} + \frac{1}{6} B_{sss}(\tilde{\xi}, \rho, 1) \tilde{s} \right] \tilde{s}^2$$

and so

$$\frac{\tilde{s}}{s} = \sqrt{\frac{\frac{e^\rho}{2} + \frac{1}{6} B_{sss}(\xi, \rho, 0) s}{\frac{e^{-\rho}}{2} + \frac{1}{6} B_{sss}(\tilde{\xi}, \rho, 1) \tilde{s}}}.$$

Taking the limit of both sides as  $s \rightarrow 0$ , we obtain

$$\frac{d\tilde{s}}{ds}(0, \rho) = \lim_{s \rightarrow 0} \frac{\tilde{s} - 0}{s - 0} = e^\rho.$$

Finally, either by repeating the above argument or observing  $\tilde{r}(\cdot, \rho)$  and  $\tilde{s}(\cdot, \rho)$  are inverse to each other, we have

$$\frac{d\tilde{r}}{dr}(0, \rho) = e^{-\rho}$$

and the second part of the lemma follows.  $\square$

Having proved the technical lemma above, we are ready to show the limit of  $D_t(B, \rho)B_\rho$  exists.

**Lemma 12.** *The function  $(s, \rho, x) \mapsto D_t(B(s, \rho, x), \rho)B_\rho(s, \rho, x)$  can be extended continuously to  $[0, 1] \times \mathbb{R} \times \partial\Delta$ . In particular,*

$$\lim_{(s, \rho) \rightarrow (0, \bar{\rho})} D_t(B(s, \rho, x), \rho) B_\rho(s, \rho, x) = 0$$

## CHAPTER 4. PROOF OF THE THEOREM

for all  $\tilde{\rho} \in \mathbb{R}$  and for  $x = 0$  or  $x = 1$ .

*Proof.* Recalling the definition of the facets  $\partial^i \Delta$  and vertices  $v_i$  of the unit simplex  $\Delta$  defined in Section (3), we have  $\partial^0 \Delta = \{1\}$ ,  $\partial^1 \Delta = \{0\}$ ,  $v_0 = 0$ , and  $v_1 = 1$ . Using (4.3.4), we have

$$\begin{aligned} D_t(B(s, \rho, v_k), \rho) B_\rho(s, \rho, v_k) &= \sum_{i=0}^1 \mu_i(\rho) h_t(B(s, \rho, v_k), \rho, \partial^i \Delta) B_\rho(s, \rho, v_k). \\ &= \sum_{i=0}^1 \mu_i(\rho) h_t(B(s, \rho, v_k), \rho, v_{1-i}) B_\rho(s, \rho, v_k). \end{aligned}$$

If  $1 - i = k$ , then we have

$$\begin{aligned} h_t(B(s, \rho, v_k), \rho, v_{1-i}) B_\rho(s, \rho, v_k) &= h_t(B(s, \rho, v_k), \rho, v_k) B_\rho(s, \rho, v_k) \\ &= -h_\rho(B(s, \rho, v_k), \rho, v_k). \end{aligned}$$

If  $1 - i \neq k$ , then  $i = k$  and we have

$$\begin{aligned} h_t(B(s, \rho, v_k), \rho, v_{1-i}) B_\rho(s, \rho, v_k) &= \frac{1}{B_s\left(h(B(s, \rho, v_k), \rho, v_{1-i}), \rho, v_{1-i}\right)} B_\rho(s, \rho, v_k) \\ &= \frac{-B_s(s, \rho, v_k) h_\rho(B(s, \rho, v_k), \rho, v_k)}{B_s\left(h(B(\tilde{s}, \rho, v_{1-k}), \rho, v_{1-i}), \rho, v_{1-i}\right)} \\ &= -\frac{B_s(s, \rho, v_k)}{B_s(\tilde{s}, \rho, v_{1-k})} h_\rho(B(s, \rho, v_k), \rho, v_k) \\ &= -\frac{B_s(s, \rho, v_k)}{B_s(\tilde{s}, \rho, v_{1-k})} \frac{\frac{1}{s}}{\frac{1}{\tilde{s}}} \frac{1}{\tilde{s}} h_\rho(B(s, \rho, v_k), \rho, v_k) \\ &= -\frac{B_{ss}(\xi, \rho, v_k)}{B_{ss}(\eta, \rho, v_{1-k})} \frac{1}{\frac{d\tilde{s}}{ds}(\zeta, \rho)} h_\rho(B(s, \rho, v_k), \rho, v_k), \end{aligned}$$

where the second equality is given by Lemma 11, and in the last equality we have

## CHAPTER 4. PROOF OF THE THEOREM

$\xi, \zeta \in (0, s)$  and  $\eta \in (0, \tilde{s})$  by three separate applications of the mean value theorem.

We conclude that  $h_t(B(s, \rho, v_k), \rho, v_{1-i})B_\rho(s, \rho, v_k)$  goes to 0 as  $(s, \rho)$  goes to  $(0, \tilde{\rho})$  for any  $\tilde{\rho} \in \mathbb{R}$  regardless of whether  $i = k$  or  $1 - i = k$  and hence  $D_t(B, \rho)B_\rho$  goes to 0 as  $(s, \rho)$  goes to  $(0, \tilde{\rho})$ .  $\square$

Thus, we have shown that (4.3.2) is continuous on  $[0, 1] \times \mathbb{R}^m \times \partial\Delta$  for all  $m \geq 1$  and we may apply Lebesgue's dominated convergence theorem to (4.3.3). Hence, the function (4.3.2) has a continuous partial derivative in  $\rho_q$  for all  $q$ . Another application of Lebesgue's dominated convergence theorem shows that the first summand (4.3.1) has a continuous partial derivative in  $\rho_q$  for all  $q$  and is given by

$$\frac{\partial}{\partial \rho_q} \int_0^\infty n[1 - D]^{n-1} S dt = \int_0^\infty \frac{\partial}{\partial \rho_q} [n[1 - D]^{n-1} S] dt.$$

### 4.3.3 Summary for $G$

Since both the first and second summand of  $G_{\rho_p}$  have continuous partial derivatives in  $\rho_q$  for all  $q$ , the function  $G_{\rho_p}$  has a continuous partial derivative in  $\rho_q$  for all  $q$ . Since we could also differentiate under the integral sign of both summands, the function  $G$  has continuous second partial derivatives and

$$\frac{\partial}{\partial \rho_q} G_{\rho_p} = \int_0^\infty \frac{\partial}{\partial \rho_q} [n[1 - D]^{n-1} D_{\rho_p}] dt,$$

for all  $p, q$ .

## 4.4 Conclusion of the proof

With all the proper pieces of the proof in place, we have proven Theorem 4.

Briefly, we summarize what we have done.

At the beginning, we started with the function

$$F_n(\rho) = \int_{\Delta^n} \max_{j=1,\dots,n} \left[ \langle \rho, \lambda^j \rangle - \langle \widehat{\lambda}^j, \log \widehat{\lambda}^j \rangle \right] d\lambda^1 \dots d\lambda^n$$

and rewrote it in Section 2.2 as

$$F_n(\rho) = \log(1 + |e^\rho|) - \int_0^\infty [1 - D(t, \rho)]^n dt.$$

We then denoted the second term, without the minus sign, by  $G_n$ . Working only with fixed  $n$ , we dropped the subscript  $n$  and simply wrote  $G$ . By Lemma 9, we concluded that both  $D$  and  $G$  are continuously differentiable in  $\rho$ . Furthermore, we could bring the derivative under the integral sign so that

$$\frac{\partial}{\partial \rho_p} G = \int_0^\infty n[1 - D(t, \rho)]^{n-1} D_{\rho_p}(t, \rho) dt$$

for any  $p = 1, \dots, m$ . In Section 4.3, we separated the integral on the right-hand side into two terms and showed that each term is continuously differentiable. Hence, we concluded the function  $G$  is twice continuously differentiable. Since  $\log(1 + |e^\rho|)$  is also  $\mathcal{C}^2$ , we conclude that  $F_n$  is twice continuously differentiable and the remaining statements of the theorem follow directly.

# Bibliography

- [BBS07] D. J. Bates, F. Bihan, and F. Sottile. Bounds on the number of real solutions to polynomial equations. *International Mathematics Research Notices*, 2007, 2007, <http://imrn.oxfordjournals.org/content/2007/rnm114.full.pdf+html>.
- [Ber75] D. Bernshtein. The number of roots of a system of equations. *Functional Analysis and Its Applications*, 9(3):183–185, 1975.
- [BRS08] F. Bihan, J. M. Rojas, and F. Sottile. On the sharpness of fewnomial bounds and the number of components of fewnomial hypersurfaces. In A. Dickenson, F.-O. Schreyer, and A. Sommese, editors, *Algorithms in Algebraic Geometry*, volume 146 of *The IMA Volumes in Mathematics and its Applications*, pages 15–20. Springer New York, 2008.
- [BS07] F. Bihan and F. Sottile. New fewnomial upper bounds from gale dual polynomial systems. *Moscow Mathematical Journal*, 7(3):387–407, July-September 2007.



## BIBLIOGRAPHY

- [BS11] F. Bihan and F. Sottile. Fewnomial bounds for completely mixed polynomial systems. *Advances in Geometry*, 11(3):541 – 556, 2011.
- [Des54] R. Descartes. *The geometry of Ren Descartes : with a facsimile of the first edition, 1637*. Dover Publications, New York, N.Y., 1954. translated from the French and Latin by David Eugene Smith and Marcia L. Latham.; : diags. ; 22 cm; Includes index.; The text (in facsimile) and translation on opposite pages.
- [Fla73] H. Flanders. Differentiation under the integral sign. *The American Mathematical Monthly*, 80(6):615–627, 1973.
- [GH94] P. Griffiths and J. Harris. *Principles of Algebraic Geometry*. John Wiley & Sons, 1994.
- [Kho91] A. G. Khovanskii. *Fewnomials ; [translated from the Russian by Smilka Zdravkovska]*, volume 88. American Mathematical Society, Providence, R.I., 1991.: ill. ; 26 cm; Translation of: Malochleny.; Includes bibliographical references (p. 133-136) and index.; Translations of mathematical monographs ; v. 88.
- [KK12] K. Kaveh and A. Khovanskii. Algebraic equations and convex bodies. In *Perspectives in Analysis, Geometry, and Topology*, pages 263–282. Springer, 2012.

## BIBLIOGRAPHY

- [MR02] G. Malajovich and J. M. Rojas. Polynomial systems and the momentum map. In *Foundations of computational mathematics (Hong Kong, 2000)*, pages 251–266. World Sci. Publ., River Edge, NJ, 2002.
- [SZ99] B. Shiffman and S. Zelditch. Distribution of zeros of random and quantum chaotic sections of positive line bundles. *Communications in Mathematical Physics*, 200(3):661–683, 1999.
- [SZ04] B. Shiffman and S. Zelditch. Random polynomial with prescribed newton polytope. *Journal of the American Mathematical Society*, 17:49–108, 2004.
- [SZ08] B. Shiffman and S. Zelditch. Number variance of random zeros on complex manifolds. *Geometric and Functional Analysis*, 18(4):1422–1475, 2008.
- [SZ11] B. Shiffman and S. Zelditch. Random complex fewnomials, I. In P. Brändén, M. Passare, and M. Putinar, editors, *Notions of Positivity and the Geometry of Polynomials*, Trends in Mathematics, pages 375–400. Springer Basel, 2011.

# Vita

Timothy Tran was born on November 25, 1985 in Fountain Valley, California. He received his Bachelor of Science in Mathematics from the University of California, Irvine in June 2008 and was accepted into a Doctoral Program at Johns Hopkins University the same year. He received his Master of Arts in Mathematics from Johns Hopkins University in May 2009. His dissertation was completed under the guidance of Dr. Bernard Shiffman, and this dissertation was defended on March 5th, 2014.